[illegible]

REVIEW AND ANALYSIS OF PERFORMANCE APPRAISAL PROCESSES

VOLUME I

[illegible]

[illegible]

[illegible]

Ministry
of
Education

86-03367/

A REVIEW AND ANALYSIS OF PERFORMANCE APPRAISAL PROCESSES

VOLUME I

A Review of the Literature

Professionalism in Schools Series

D. A. ONDRACK, Principal Investigator

C. OLIVER, Research Associate

This research project was funded under contract by the Ministry of Education, Ontario.

This study reflects the views of the authors and not necessarily those of the Ministry of Education.

The Honourable Sean Conway, Minister

© QUEEN'S PRINTER FOR ONTARIO, 1986

Order Information:

Publications Sales
The Ontario Institute
for Studies in Education
252 Bloor Street West
Toronto, Ontario M5S 1V6
Will invoice on orders over $30.00.
Other orders must be accompanied by a
cheque or money order payable to
O.I.S.E.

Publications Services
880 Bay Street, 5th Floor
Toronto, Ontario M7A 1N8

(416) 965-6015
(Toll Free) 1-800-268-7540
(Toll Free from area code 807)
Ask operator for Zenith 67200
Order must be accompanied by a cheque
or money order payable to the
Treasurer of Ontario

ONO 1823

Canadian Cataloguing in Publication Data

Ondrack, D. A.
A review and analysis of performance appraisal
processes

Includes bibliographies.
Partial contents: v. 2. Performance appraisal for
professional service employees, field study
report -- v. 3. Performance appraisal for
professional service employees, non-technical
report.
ISBN 0-7729-1087-1 (3 v. set). -- 0-7729-1088-X
(v. 1). -- 0-7729-1089-8 (v. 2). -- 0-7729-1090-1
(v. 3)

1. Teachers--Rating of. 2. Employees, Rating of.
I. Oliver, Christine Tomaszuk. II. Ontario.
Ministry of Education. III. Title. IV. Title:
Performance appraisal for professional service
employees.

LB2838.O52 1986 371.1'44 C86-099627-1

TABLE OF CONTENTS

Exhibits and Figures vi

Acknowledgements vii

Abstract viii

Introduction to the Study 1

A. Performance Appraisal in Context 6
  1.0 Professional Bureaucracies Theory 8
  2.0 Performance Appraisal Overview 10
  3.0 Environmental Trends 14
    3.1 Legal Environment 14
    3.2 Economic Environment 15
    3.3 Social Environment 17
    3.4 Demographic Trends 18
  References 21

B. Performance Appraisal: General Review 24
  4.0 Purposes of Appraisal 24
    4.1 Monitoring and Control 24
    4.2 Feedback and Development 25
    4.3 Compensation Administration 25
    4.4 Promotion/Retention/Tenure/Transfers 26
    4.5 Human Resources and Program Planning 26
    4.6 Organizational Accountability 27
  5.0 Design of Appraisal Systems 28
    5.1 Process vs. Output-Focused Systems 28
    5.2 Hierarchical Differences in Appraisal Systems 32
  6.0 Development of Performance Standards 33
    6.1 Characteristics of Effective Standards 33
    6.2 Personality, Skills and Behaviour Criteria 35
      6.2.1 Personality Criteria 35
      6.2.2 Skills Criteria 36
      6.2.3 Behaviour Criteria 38
      6.2.4 Results Criteria 38
  7.0 Appraisal Techniques 40
    7.1 Classification of Techniques 40
    7.2 Individual or Absolute Appraisal Techniques 41
      7.2.1 Essay Evaluations 41
      7.2.2 Graphic Rating Scales 43
      7.2.3 Behaviour Checklists 44

7.2.4 Forced-Choice 45
7.2.5 Critical incidents 46
7.3 Multiple Person or Relative Appraisal Techniques 47
7.3.1 Ranking 47
7.3.2 Paired Comparisons 47
7.3.3 Forced Distribution 48
7.4 Other Appraisal Techniques 48
7.4.1 Behaviourally Anchored Rating Scales (BARS) 48
7.4.2 Management by Objectives 50
7.4.3 Assessment Centres 51
8.0 Implementation of the System 52
8.1 Choice of Appraiser 52
8.1.1 Supervisory Appraisal 53
8.1.2 Peer Appraisal 54
8.1.3 Appraisal by Subordinates 54
8.1.4 Self-Appraisal 55
8.1.5 Appraisal by Outsiders 56
8.2 Appraiser Training 57
8.3 Assessment Period and Timing 59
8.4 Appraisal Interviews 60
8.5 Linkages of Appraisal Systems to Other Personnel Systems 61
8.5.1 Career Planning and Counselling 61
8.5.2 Human Resource Planning and Inventory Systems 65
8.5.3 Promotion, Tenure and Dehiring 65
8.5.4 Compensation Administration 66
8.5.5 Training and Development 67
9.0 Results 67
9.1 Organizational Performance and Employee Satisfaction 67
9.2 Maintenance of the System 70
10.0 Conclusion 72
References 75

C. Teacher Evaluation Literature Review 87
11.0 Need for an Evaluation System 87
12.0 Critical Issues and Components 91
12.1 The Criteria Question 91
12.1.1 Presage Criteria 91
12.1.2 Output-oriented Criteria 92
12.1.3 Process-oriented Criteria 93
12.1.4 Discussion of Criteria 96

12.2 Choice of Appraiser 97
    12.2.1 Teacher Self-Evaluation 97
    12.2.2 Multiple Appraisers 98
    12.2.3 Peer Appraisal 98
    12.2.4 Appraisal by Outside Experts 99
    12.2.5 Appraisal by Students 99
    12.2.6 Discussion of Appraisers 101
12.3 The Role of Goal Setting 101
12.4 Formative and Summative Evaluation 104
12.5 Components of a Good Evaluation System 107
13.0 Constraints on the Evaluation Process 109
13.1 Lack of Validated Research 109
13.2 Teacher Resistance 110
13.3 Criteria Ambiguity 111
13.4 Teaching as a Contextual Experience 113
13.5 The Teacher as Professional 114
13.6 Evaluator Resistance and Ability 115
14.0 Approaches to Teacher Evaluation 116
14.1 Current Conventions 116
14.2 Variations to Conventional Approaches 120
14.3 Innovations in Teacher Evaluation 128
    14.3.1 Teacher Testing 128
    14.3.2 Performance Objectives Approach 129
    14.3.3 Behaviorally Anchored Rating Scales 133
15.0 Summary 134
16.0 Conclusion 136

Appendix 148
References 156

Exhibits and Figures

Exhibit B-1: Appraisal System Design: Advantages and Disadvantages 31

Exhibit B-2: Criteria for Choice of Performance Evaluation Techniques 42

Exhibit B-3: Guidelines for Appraisal Interviews 62

Figure B-1: Linkages of Appraisal Systems to Other Personnel Systems 64

Figure C-1: The Link Between Evaluation and The Quality of Education 88

Exhibit C-1: Example of an Inventory of Competency Statements 95

Figure C-2: Framework for Establishing Appraisal Standards 112

Exhibit C-2: Principles, Rules, and Errors to be Avoided in Classroom Observation 118

Exhibit C-3: Common Classroom Assessment Instruments: A Summary 122

Exhibit C-4: Sample Rating Scales 125

Exhibit C-5: Performance Objectives Approach 130

Exhibit C-6: Multifaceted Appraisal System: A Hypothetical Example 141

ACKNOWLEDGEMENTS

The authors of this study wish to acknowledge the significant contributions of all field site personnel whom we interviewed for our study. Without their assistance and cooperation this project would not have been possible. Special recognition should be given to Dr. Glen Tecker whom we interviewed at Educational Testing Services, Princeton, New Jersey. His insights into the process and methods of appraisal were both inspiring and highly informative.

The authors would also like to gratefully acknowledge the guidance and assistance of Dr. Peter Nightingale of the Ministry of Education over the two years of work involved in this project. His enduring patience and ability to act as an effective liaison between all the advisors and participants in the various related research projects were crucial to the progress and completion of the project. We are also very appreciative of the comments and assistance received from the members of the Advisory Committee, particularly Noel Clark, Bruce Archer, and Michael Prokopich. They have been involved in the Advisory Committee from the beginning and we strongly suspect that their involvement will continue in subsequent developments.

Abstract

This research project consisted of two parts: a review of the literature on performance appraisal with special attention to the literature on certificated personnel in public education systems (Vol. 1), and field site studies of performance appraisal in action at ten professional service organizations in Ontario and eight public school systems in the U.S. (Vol. 2). Each of the sites studied was selected on the basis of reputation for advanced practices in appraisals.

The literature review showed that there were two predominant approaches to conducting performance appraisals. One focuses on the "process" of professional work, or the techniques, means, skills, and styles used to perform professional service. The other focuses on the "outputs" of professional work, or various measures, outcomes, or judgements of the results of professional work. As well, the literature review showed that there were two main uses of appraisals: in "formative" uses, appraisal information is used for recognition, guidance, coaching, development, or problem correction in professional work; in "summative" uses, appraisal information is used for administrative decisions on personnel such as compensation, promotion, training, transfers, or discipline. These approaches to, and uses of performance appraisal are not mutually exclusive and various combinations can be found in many systems.

Finally, another major theme in the literature review was the growing emphasis on protection of employee rights legislation and the vulnerability of employers to lawsuits arising from contentious personnel decisions based on performance appraisals. These trends have led to considerable effort to improve the reliability and validity of appraisal systems, but have also increased the bureaucratization of procedures.

The field site studies mirrored many of the results from the literature review with three important additions:

1. Many appraisal systems are in constant evolution so that what is reported in the literature is often not in practice when the site is visited.

2. Different approaches to and priorities in appraisal are used depending upon the local situation of the organization.

3. The bureaucratic cumbersomeness of many appraisal systems causes short-cuts to be used. In practice, this means that most teachers/professionals receive at least "satisfactory" ratings and the elaborate quasi-judicial safeguards built into many systems are used only rarely.

While no generally suitable appraisal system was found by means of either the literature review or the field site study, two principles seemed to operate in the most satisfactory sites:

a) Under a set of general principles or guidelines established by a relevant central authority, a decentralized appraisal system is developed in operational detail by the local organization to reflect local concerns and priorities.

b) The decentralized system is developed jointly by supervisors and subordinates in the organization so that all involved parties accept the relevance and legitimacy of the appraisal system. Locally developed systems which satisfy overall standards set by the larger organization seem to be most successful at working through the inevitable learning curve of implementation and adaptation of an effective appraisal system.

Introduction to the Study

This contract research project was authorized by the Ministry of Education of Ontario as one of a series of research studies on performance appraisal in the public education system. The purposes of this study were to conduct a review and analysis of published literature on performance appraisal in general and in particular on the use of appraisals in public education systems. Volume 1 is the literature review. As well, the study consisted of a series of field site investigations of performance appraisal systems in action to compare actual practice with results reported in the literature which is volume 2. Two types of field sites were chosen for the study: a sample of professional service organizations in Ontario which employed a large number of professional service employees, such as researchers, consultants, and educators; and a sample of public school boards in the United States which had a reputation for advanced practices in the appraisal of teachers. The results from the field studies provided an opportunity to compare information from the literature with information on appraisal practices in organizations which might be analogous to public school boards in Ontario. In this way, the vast range of information on performance appraisal could be focused and made relevant to the needs of the public education system in Ontario.

In summary, the specific objectives of the study were:

* to review the state-of-the art theory and practice of performance appraisal approaches and methodologies;

* to identify those purposes of performance appraisal that might be appropriate to an education system;

* to identify those approaches which might be particularly relevant to the needs and practices of Ontario's educational system.

The results of this study are intended to provide a common base of information for use by the parties in Ontario most responsible for, and most affected by, performance appraisal of certificated education staff. To insure that all parties shared the same perception of the nature of the study and the same access to its results, an Advisory Committee was established to oversee the definition and progress of the study. The Advisory Committee was composed of representatives from the Ministry, from trustee groups, and from professional associations involved in public education in Ontario, as well as representatives from other related contract research projects. Progress reports by all the contractors were presented periodically and discussed in Advisory Committee meetings to ensure coordination between projects and to obtain advice and commentary from the committee.

In both the review of the literature and the field site studies, the report will focus on the following subjects:

* Purposes of performance appraisal -- This subject concerns how appraisals are used or what purposes are served by appraisals in the organization. Some examples of purposes are personnel planning, accountability, staff growth and development, compensation and dismissal decisions, quality control, motivation, and achievement of organization objectives.

* Performance criteria -- This subject concerns the development of relevant criteria for different aspects of performance to provide fair and accurate standards with which it may be assessed. Such issues as the reliability, validity, and acceptability of criteria are discussed in this section.

* Performance data collection -- Once criteria have been established, data must be gathered according to the criteria. This section is concerned with the various means by which data are gathered and how the data are organized so as to be useable in the appraisal process.

* Appraisal methods -- This section is concerned with an exploration of the various techniques used to conduct appraisals and to form judgements about a person's work for both process and output appraisal systems.

* Appraisal logistics -- This topic is closely related to appraisal methods, but focuses on the questions of who conducts appraisals and how often, as well as on the conditions under which the logistics might be varied. For example, the appraisal methods and logistics used for probationary employees would differ from those used for tenured employees.

* Appraisal training -- In a broad sense, this subject is concerned with the means by which participants in the appraisal process learn to implement and maintain the system. Training normally occurs for supervisors in an organization, but could also

occur for subordinates as well. Related to training are organization support services for the appraisal process, such as communications bulletins, procedural manuals, and monitoring activities for quality control of the appraisal process.

- Evaluation and maintenance of the system -- This subject is concerned with the determination of the impact of various approaches to appraisal on the organization and on individual performance. Maintenance is concerned with the means by which appraisal systems are monitored and their effectiveness maintained. This is a difficult and complex aspect of appraisal systems, but it is crucial to the ongoing adaptability and relevance of an appraisal system.

- Trends and influences on appraisal systems -- Closely related to the issue of monitoring and adaptation is the issue of the responsiveness of an appraisal system to the impact of changes in the environment of the organization. This involves an examination of significant trends in the legislative, political, and economic environments as they pertain to the present and future viability of various appraisal systems.

Results of the reviews of the general management literature on appraisal and teacher evaluation literature were organized according to the above subject. In this way a systematic comparison could be done between the two sets of literature to see which parts or aspects were unique to education organizations and which parts seemed to be generally applicable to all organizations. A similar type of systematic comparison was also done with the results of the field site organizations. The field site organizations were regarded as being analogous in various degrees to Ontario public school system organizations. All of the organizations were similar in that they all employed a high proportion of professional service employees engaged in client service work. Some degree of measurable outputs from the professional service could be obtained in all cases, but typically, many intangible aspects of performance were also associated with the various types of professional service.

The field sites were grouped into three subgroups as follows:

i) a Business group composed of five-profit based or self-financing organizations providing professional services such as research, consulting, and teaching;

ii) a College group composed of five colleges of applied arts and technology (CAATs) in Ontario which are postsecondary, tuition-based educational organizations;

iii) a U.S. schools group composed of seven leading public school systems and one training and consulting organization for teacher evaluation.

All of these organizations are similar to Ontario public school systems in various ways and of course are also different in various ways. These aspects of the field site sample will be discussed in greater detail in volume 2 the section which concentrates on the field site portion of the research project. In general, the study of analogous organizations focused upon the following issues:

- identification of the appraisal methodologies and the uses of appraisals in the various types of organizations;
- identification of major differences (if any) in approaches to appraisal which seem to be required by the different organizational types;
- identification of differences in approaches to appraisal in different functions of the organization and at different levels in the hierarchy;
- identification of whether process or output approaches to appraisal are typically used in some functions and organizations and not in others;
- identification of various approaches to developing criteria and gathering data on the work of professionals;
- identification of linkages between the appraisal process and other aspects of human resource management in the organization, such as compensation systems, training and development, career planning, and supervision.

Volume 3 of the report provides an integration of the results from the various investigations as well as the authors' conclusions. Since this study did not look directly at public education systems in Ontario (this was done in other related research projects), specific recommendations and conclusions for Ontario are not given. Instead the authors have, as required, provided a common information base for participants in the development of policy for the education system in Ontario and it will be these persons who eventually develop specific conclusions and recommendations. This report provides, to the best of our judgement, a set of generally applicable conclusions about an effective system of performance appraisal in education, and it is hoped that these will establish a useful foundation of guidelines for future development.

A. PERFORMANCE APPRAISAL IN CONTEXT

While it is true that the issue of performance appraisal in both the private and public sector has come under increasingly greater scrutiny in recent years, surveys[1] reveal that there has been no concomitant increase in the level of satisfaction with currently applied systems. Nor has greater attention to the subject led to a convergence of wisdom on its design, use, and implementation. If the complexities of performance appraisal application seem too numerous to unravel, why do private and public organizations continue to persevere? The answer lies in a number of factors:

* Inevitably, organizations attempt to determine the extent to which their employees are performing their jobs effectively, either formally or informally; therefore it seems wise to systematize the process.
* A wide variety of important uses and positive results are associated with or served by well-developed appraisal programs.
* There is a rapidly growing movement toward accountability as the means by which an organization maintains, develops, and satisfies its personnel.
* There is a strong "justificatory" trend developing in all business and public service fields; that is, a need to justify and defend organizational decisions and the quality of the organization's output.

Therefore, despite a considerable amount of disenchantment with the entire performance evaluation process, there remains an obvious need for the continued institution or revision of formal systems of performance appraisal.

The complexities of objective performance evaluation take on added significance in the realm of education. For here improvement in the quality of instruction is a necessary but not sufficient outcome of successful evaluation. In the face of a growing accountability movement, evidence of improvement may be at least as important. Unlike in the past where students were held primarily accountable for their performance in school, teachers and school administrators are now expected to accept increased responsibility for the realization of objectives relating to changes in students' achievement or behaviour. This places extra demands on the effectiveness of performance appraisal in the educational context because evidence of improvement in the quality of instruction may not be easily forthcoming. There are two particular reasons why this is the case. First, knowledge of cause and effect relations in teaching is incomplete due to the complex relations among a number of potential causal factors (the influence of a student's home environment, for example); hence it is difficult to attribute outcomes to the specific acts or behaviours of educational personnel. Second, the appropriate standards for assessing performance are ambiguous and controversial.

---

[1] See James W. Steele, "In Praise of Appraisal", S.A.M. Advanced Management Journal, Summer 1981, p.12; or Alan H. Locher and Kenneth S. Teel, "Performance Appraisal -- A Survey of Current Practices", Personnel Journal, May 1977, pp. 245 - 247, 254.

This suggests that when direct evidence of instructional improvement cannot be fully ascertained, faith in the process of performance appraisal is paramount. (Consider, for example, our judicial system: although the outcome of "guilty" or "not guilty" is difficult to assert unequivocally, we tend to trust the decision-making process by which that outcome is achieved). Accordingly, performance appraisal processes in an education system must not only be validly designed, but also acceptable to both users of the system and to the public. In other words, performance appraisal applicable to education staff must indirectly satisfy external constituents (parents, the community) at the same time that it serves the needs of its internal members for purposes of feedback, tenure and redundancy decisions, monitoring the implementation of policies, programs, and professional growth or development.

The next section reviews Henry Mintzberg's theory of Professional Bureaucracies, in order to place the special needs and opportunities of the educational system in context for subsequent consideration of appraisal issues. This will be followed by an overview of performance appraisal and an analysis of environmental trends relevant to the application of performance appraisal generally and within the realm of education particularly.

1.0 Professional Bureaucracies Theory

In his work entitled The Structuring of Organizations,[2] Henry Mintzberg has classified organization structures into different typologies which display certain distinctive features. The "Professional Bureaucracy", common in universities, general hospitals, social work agencies, public accounting firms, and school systems, is a type of structure which "hires duly trained and indoctrinated specialists -- professionals -- for the operating core, and then gives them considerable control over their own work".[3] About the teacher, it is stated that: "The teacher works alone within the classroom, relatively hidden from colleagues and superiors, so that he has broad discretionary jurisdiction within the boundaries of the classroom."[4]

There are several characteristics which are features of the Professional Bureaucracy. Initial training for professionals, often an extended or rigorous affair, ensures a standardization of knowledge and skills so that each professional can be expected to work relatively autonomously according to certain predetermined criteria. Nevertheless, considerable discretion remains in the application of one's duties. Furthermore, as new knowledge is generated and new skills develop, the professional upgrades his or her expertise.

The Professional Bureaucracy is a highly decentralized structure, in which professionals tend to identify more with their profession than with the organization in which they work. In addition, "not only do the professionals control their own work, but they also seek collective control of the administrative decisions that affect them."[5] What then, is the role of the professional administrator? There are two primary tasks: handling disturbances in the organization and serving "key roles at the boundary of the organization between the professionals inside and interested parties -- governments, client associations and so on --

---

[2] Henry Mintzberg, The Structuring of Organizations (Englewood Cliffs, N.J.: Prentice-Hall Inc., 1977).

[3] Ibid., p.349.

[4] Ibid.

[5] Ibid., p.358.

on the outside".[6] This latter role might include securing important resources for the professional or buffering the individual against external pressures.

The power of the professional resides within his or her own expertise. Given the autonomy of the professional and the nature of his or her work, Mintzberg states that "the work processes themselves are too complex to be standardized directly by analysts ... the outputs of professional work cannot easily be measured and so do not lend themselves to standardization ... thus, Professional Bureaucracies cannot rely extensively on the formation of professional work or on systems to plan and control it."[7]

The implications for the evaluation of performance in education are obvious. If criteria are to be output-based, the difficulty of measuring the outputs of professional work makes the development of standards against which to measure performance a formidable task. (Hence the controversy surrounding "pupil progress" as a criterion). Furthermore, Mintzberg states that "specific professional activities -- complex in education and vague in results -- are difficult to control by anyone other than the professionals themselves."[8] At the same time, however, it is pointed out that "the professional bureaucratic structure cannot easily deal with professionals who are either incompetent or unconscientious."[9] Therefore, while it is evident that some form of control is required to ensure staff competence, the nature of teaching does not easily lend itself to direct supervision.

The Professional Bureaucracy as depicted by Mintzberg serves as a framework for the consideration of challenges inherent in designing an effective performance appraisal process for professional education staff. It is not surprising, in light of this theory, that the area of teacher evaluation has been historically fraught with conflict and controversy. The same difficulties, however, need not apply at the director, superintendent or principal level for this reason: performance appraisals for certificated education staff probably should not focus on a single approach at all levels of the hierarchy. While school supervisory policy needs to be integrated at all levels, appraisal and supervision must be tailored to the goals and tasks of different hierarchical levels. What is needed then, in the educational field is a "system of appraisals", rather than an "appraisal system".

---

[6] Ibid., p.362.

[7] Ibid., p.352.

[8] Ibid., p.376.

[9] Ibid., p.373.

Professional Bureaucracy theory is not prescriptive in nature; rather, it provides a context within which to consider the design and implementation of performance appraisal among teachers and emphasizes the need to differentiate the special concerns of the appraisal of teaching from appraisal at higher levels of the educational system. These issues will be elaborated in Part C of this study.

2.0 Performance Appraisal Overview

The search for a definition of "performance appraisal" leads one to ask the fundamental question: What does an organization want to accomplish through implementation of a performance appraisal system? Consider the following definitions:

* Performance appraisal is the systematic evaluation of the individual with respect to his performance on the job and his potential for development.[10]

* Performance appraisal or evaluation is the process of identifying, measuring and developing human performance in organizations.[11]

It is clear that the essence of performance appraisal is performance measurement. Further, the measurement of performance is assumed to be formal or relatively systematic. Why does an organization wish to formally measure performance?

Organizations typically seek to achieve two purposes when they appraise performance. The first of these pertains to the valuative or judgemental function of performance appraisal. This function, which concentrates on past and current performance, forms the basis for administrative decisions relating to promotion, salary increases, retention, dismissal, discipline, and human resource planning. The second major use of appraisals is developmental. In this regard, appraisals focus on improvement in future performance and the potential for employee growth by means of coaching, counselling, goal setting, and attempting to motivate the employee to be more productive, effective, efficient and/or satisfied.

---

[10] Dale S. Beach, Personnel: The Management of People at Work (New York: Macmillan Publishing Co. Inc., 1980), p.290.

[11] Lloyd Baird, Richard W. Beatty, and Craig Eric Schneier, eds., The Performance Appraisal Sourcebook (Amherst, Mass.: Human Resource Development Press, 1982), p.4.

Historically, performance appraisal systems have emphasized the evaluative function. Although current practices vary widely, there is a movement toward incorporating more and more of the developmental functions into the appraisal process. Combining the evaluative and developmental functions within a single appraisal approach tends to place the appraiser in the paradoxical role of judge and counsellor, a role with which most managers or supervisors feel uncomfortable. Solutions to this paradox are divided among four different schools of thought. First, it has been suggested, for example, that "many performance appraisal programs encounter difficulty in attempting to serve more than one purpose ... the most effective programs are those focused on a single, clear defined purpose."[12] Other researchers[13] purport that it is sufficient to "uncouple evaluation and development" by means of two separate performance appraisal interviews. A third approach is the use of two separate interviewers. Finally, it is sometimes held that the ambivalence created by the supervisor's dual role is best ameliorated by proper training to ensure that the two roles can be handled simultaneously. Whichever approach is taken, the importance of identifying the intended use, functions, and objectives of performance appraisal cannot be overemphasized. An organization using a performance appraisal system must be able to respond unequivocally to the question: For what reason(s) do we wish to systematically measure performance?

A study conducted by the Conference Board in 1977 highlighted a number of problems in performance evaluation practices. In its state-of-the-art review of how nearly 300 U.S. corporations regard performance appraisals, the study concluded in part that:

* formal performance appraisals are not so much accepted as they are tolerated;

---

[12] John D. McMillan and Hoyt W. Doyel, "Performance Appraisal: Match the Tool to the Task", Personnel, July-August 1980, p. 13.

[13] See, for example, Michael Beer, "Performance Appraisal: Dilemmas and Possibilities", Organizational Dynamics, Winter 1981, p.29.

- current systems are widely regarded as a nuisance at best and a dangerous evil at worst;
- a main concern of critics is a belief that present systems may generate unreliable or invalid data.[14]

This study serves as evidence of the challenges ensuant to system design.

First, data generated from formal performance measures must be valid and reliable. In other words, the appraisal process must yield consistent data, regardless of who does the appraisal or when it is conducted. Furthermore, there must be a significant correlation between the work behaviours being measured and those which are critical requirements of the job.[15] Birkerhoff and Kanter (1980) suggest that this is most likely to occur when:

- the purpose of the appraisal is clear
- the tasks are simple
- goals for the tasks are clear
- outcomes are predictable
- tasks are relatively independent
- task performance is observable
- criteria for performance are set by those later assessing performance
- appraisers feel secure in their own jobs and have no personal stake in hurting the performer.[16]

A second prerequisite of good system design is its acceptability to users of the system. As Dipboye and Pontbriand (1981) state, "employee opinions regarding the appraisal process may be as crucial to its long-term effectiveness as the validity and reliability of the measures."[17] Third, the evaluation process must be practical and workable. A frequent complaint regarding the performance appraisal process is the administrative burden and length of time required to conduct the appraisal. Fourth, the approach must be standardized in its design and consistent in its administration. Finally, the appraisal program must be based on careful job analysis to ensure job-relatedness.

---

[14] See Selig M. Danzig, "What We Need to Know About Performance Appraisals", Management Review, February 1980, p.20.

[15] See George F. Truell, Performance Appraisal: Current Issues and New Directions (New York: PAT Publications, 1980), p.14.

[16] Derick W. Brinkerhoff and Rosabeth Moss Kanter, "Appraising the Performance of Performance Appraisal", Sloan Management Review, Spring 1980, p.13.

One remaining component in design relates to the issue who participates in designing the process. Generally speaking, involvement in the design process increases commitment to the system. Therefore, where it is feasible, those who will use the system should be encouraged to participate in the design phase of the program.

The substratum of valid, reliable and acceptable appraisal design is the development of accurate criteria against which to measure performance. These criteria may be based on the outputs or results of work, the process or series of tasks by which results are achieved, or on a combination of the two. The performance of the employee may be described or assessed in terms of his or her behaviour, perceived personality traits, or accumulated repertoire of skills and knowledge. The choice of design features among these alternatives should be contingent upon the nature of the employee's task and work environment.

The first step in implementation involves an assurance that the evaluation of individuals is compatible with, and contributes to, the goals and policies of the overall system or organization. In an educational system, for example, program objectives must be compatible with instructional objectives and evaluations must be coordinated from the classroom level up to ensure improved school operations and consistency with overall district policy. Similarly, the introduction of effective procedures for enhancing accountability at the individual level should be a reflection of the need for accountability throughout all levels of the system.

---

[17] Robert L. Dipboye and Rene Pontbriand, "Correlates of Employee Reactions to Performance Appraisals and Appraisal Systems", Journal of Applied Psychology, April 1981, p.248.

Implementation of the performance appraisal system relates to the actual logistics of evaluation:

- Who will conduct the appraisal (superior, peer, subordinate, self, outsider)?
- What personnel training is required to bring the system on stream?
- How often should appraisals be conducted?
- Who should receive feedback and in what form?
- How should the interview be conducted to derive maximum benefit for the appraiser, appraisee and the organization?
- How will data obtained from the interview be utilized and to whom should that data be made available?

These and other questions will be explored further in subsequent sections of this study.

The benefits offered by a performance appraisal system can be substantial. For the employee, it can provide satisfaction, motivation, and an opportunity for growth and development. For the organization, it can improve the effectiveness of administrative decisions relating to human resource management, enhance organizational performance, and provide defensible grounds for the conduct and output of the organization system. There appears to be wide agreement that currently applied performance appraisal systems require improvement, but in both the private and public sector the need for some form of appraisal is becoming an article of faith, even among its most vocal critics.

3.0 Environmental Trends

There are a number of trends in the legal, economic and social environments of Canada and the United States which are having a catalytic effect on the search for better appraisal methods. Major changes in each of these aspects of the environment are discussed in the following sections.

3.1 Legal Environment

In the United States, the impact of equal employment opportunity (EEO) legislation and court rulings in the 1970's has mandated the widespread development of formal, validated systems of performance management. As Hersey and Goldsmith (1980) point out, "for the first time, U.S.

organizations face the prospect of having to make personnel decisions in ways that can be justified to the federal government."[18] The two most significant pronouncements on the issue are the "Uniform Federal Guidelines on Employee Selection Procedures", a 1978 attempt to unify federal selection procedures, and the 1971 U.S. Supreme Court Griggs vs. Duke Power Co. decision.[19] In view of the integral part performance appraisal may play in making selection, promotion, compensation, and dismissal decisions, a number of court decisions have ruled on whether or not particular performance appraisal methods are in compliance with civil rights legislation. The impact of this legislative pressure has been twofold: there has been an increase in the rate of adoption of performance appraisal systems to document administrative decisions for legal protection, and there has been an acceleration in the search for more objective and valid appraisal techniques.

Canada has by and large pursued a much less aggressive role in civil rights promotion. Although the underlying theory and purpose is the same, our human rights legislation, and more particularly its administration, is not as sophisticated or well-developed. This is due in part to a shorter history and lower number of judicial decisions, but the practical effect is that there is less direct intervention in an organization's employment practices. It is a matter of speculation whether the significant impact that legislation is having on performance appraisal programs in the United States will ever be mirrored in the Canadian legal system.

3.2 Economic Environment

Presently, Canada's economic environment is characterized by increasing inflationary trends, productivity slowdowns, greater government regulation and intervention, high unemployment, and a slackening of investment. In the private sector, business is being forced to take a hard-line approach to the productivity potential of human resources. As Kahalas (1980) emphasizes: "Organizations will find it necessary to alter their performance evaluation practices in order to maintain or improve their human resource utilization and overall organizational effectiveness."[20] Similarly, in the public sector, the recent downturn in Canada's economy is leading to a call for restraint in government spending and more emphasis on cost-effective deployment of government resources which in turn places an additional burden of proof on the public service sector to justify its use of resources. In both the public and private sectors, with personnel continuing to grow as a significant fixed cost in many organizations, there is more pressure to identify the good performers and remove the weaker ones. For this reason, a well-designed appraisal program becomes particularly necessary to the efficient management of human resources. However, in Canada the obvious need for these programs is not being translated into an increase in the implementation of systems comparable to that being witnessed in the United States. Ryval (1981) suggests that "much of Canadian industry is lagging anywhere from five to 10 years behind current development in PA [performance appraisal]."[21] The same claim may not necessarily be made with regard to the public sector. The government of Ontario, for example, initiated a system of performance appraisal in January 1980 for over 80,000 employees.

---

[18] Paul Hersey and Marshall Goldsmith, "The Changing Role of Performance Management", Training and Development Journal, October 1980, p. 18.

[19] For more information on American legal requirements, see Patricia Linenberger and Timothy J. Keaveny, "Performance Appraisal Standards Used by the Courts", Personnel Administrator, May 1981, pp.89-94; or Barry J. Baroni, "The Legal Ramifications of Appraisal Systems", Supervisory Management, January 1982, pp.40-43.

The main point to recognize here is that economic constraints place greater demands on organizations to justify and utilize their human resources efficiently. In order to maximize the use of resources and achieve organization objectives, whether for profit-making or public service purposes, it is necessary to evaluate the extent to which employees are performing in a manner conducive to overall goal attainment. As the trend toward increased accountability to society in both the public and private sectors continues to gain momentum, this aspect of human resource management will become more and more critical.

---

[20] Harvey Kahalas, "The Environmental Context" of Performance Evaluation and Its Effects On Current Practices", Human Resource Management, Fall 1980, p.32.

[21] Michael Ryval, "The Corporate Report Card", The Financial Post Magazine, March 1981, p.45.

3.3 Social Environment

Social pressures on performance evaluation practices stem largely from changing attitudes about work. Daniel Yankelovich, president of a large social research firm, claims that people who work are no longer motivated for the same reasons to work as hard and as effectively as in the past.[22] In the late 1960's almost half of all employed Americans looked to their work as a primary source of personal fulfillment. Now that number has plunged to fewer than one out of four.[23] At the same time, today's work force has higher occupational expectations. It may be inferred from this that as employees demand more satisfaction and opportunities for growth at the work place, the importance of the developmental function of appraisals will rise. Levinson, in his classic article "Management by Whose Objectives?", claims that appraisal systems must be designed so that the employee's personal needs, values and goals are incorporated into, and compatible with, his or her work goals.[24]

Social values are also changing. There is an increasing desire (particularly among white-collar workers) for participation in decision-making. In light of the fact that there has been a significant increase in this phenomenon, the implication to be drawn is that the "shared appraisal" may become more prevalent. In this way, the appraisee's role is likely to evolve from one in which he or she is passively evaluated to one in which he or she participates actively in the process. The demand for more rights at the workplace may, in the extreme, result in a resistance to evaluation altogether. In the area of education, the need for autonomy among professional staff is particularly strong, as the preceding outline of Professional Bureaucracy theory indicated. A recent report of the Canadian Education Association seminar held in 1978 states the following:

> The rise of professionalism among teachers has reached a point where, more and more, they are demanding the exclusive right to determine the goals of education and to choose the ways and means of achieving them. They deny any jurisdiction on the part of their superiors to evaluate the scope and the effect, good or bad, of the teacher's professional activity ...[25]

---

22. Daniel Yankelovich, "Yankelovich on Today's Workers: We Need New Motivational Tools", Industry Week, August 6, 1979, pp.2-9.

23. Kahalas, "Environmental Context", p.33.

24. Harry Levinson, "Management By Whose Objectives?", Harvard Business Review, July-August 1976, pp.125-134.

25. Canadian Education Association, The Supervision of Teaching (Report of a CEA seminar held at Montreal, July 1979), p.48.

This is an obvious indication that performance appraisal systems must offer considerable benefits to all participants in the process in order to overcome such resistance.

Another manifestation of changing social values is that employees want and demand more information about themselves. This suggests that the feedback aspect of appraisal may also become increasingly important. Furthermore, a recent opinion survey by Louis Harris and Associates found that "88% of the responding workers felt they should have the right to look at and question supervisors' reports to management on their performance and promotability."[26]

The question of performance appraisal for control purposes is complicated by changing values regarding freedom and individualization at the work place. Moreover, this is particularly true among professionals: "in the field of education, especially, there is a rejection in many milieus of any attempt to control, on the pretext that only the teacher has the ability to assess his own competence."[27] Some means of monitoring organizational and individual effectiveness must remain, however, in order to ensure that system-wide goals are achieved. Performance appraisal at the classroom level should be designed to accommodate the discretionary needs of the professional, but in such a way that data for monitoring and correctional purposes can be obtained. The nature of professional work, coupled with changing social values regarding employee rights, dictates an appraisal design which is sensitive to the needs and values of the appraisee.

3.4 Demographic Trends

Population trends are also indirectly influencing the use of performance appraisal. While the total labour force growth will halve in the decade ahead, the prime 25- to 44-year-old work force will increase at roughly double the average rate, and the level of women's participation will continue to increase in the 1980's. "The changing demographics of the labour force are

---

26. Kahalas, "Environmental Context", p.34.
27. Canadian Education Association, The Supervision of Teaching, p.48.

sharply reducing the supply of new workers that will reach organizations in the 1980's; while the supply of prime working-age employees will grow, companies will be less able to replace employees with new, young workers."[28] For organizations in general, three potential uses of performance appraisal may receive more emphasis: human resource planning, motivation of "plateaued" personnel, and identification of retraining needs. In an increasingly service-oriented economy, new positions which evolve will have to be staffed more and more by current work force members. This may entail "redundancy planning", that is, retraining incumbents in traditional jobs for new positions. Additionally, as the "bulge group" created by the postwar demographic surge ages, efforts will have to be made to counteract a loss of motivation and decline in performance among those members who remain in the same position until retirement.

In education, the influence of demographic changes and declining enrolment possess similar implications with regard to the latter factor of motivation. The lack of turnover among a "locked in" teaching force creates the need for continued renewal among staff members.

> Declining enrolments present us with harsh reality: while there will be fewer probationary teachers in our schools to absorb the time of evaluators, there will be the problem of loss of motivation, of incentive, among the remaining permanent teachers.[29]

This suggests the importance of the developmental function of performance appraisal.

Environmental factors are beginning to exert pressures on performance appraisal programs. Legislation, economic pressures, social changes and demographic trends are requiring that programs be re-examined to bring about a more mature and effective activity in this area. The

---

28. Ann M. Morrison and Mary Ellen Kranz, "The Shape of Performance Appraisal in the Coming Decade", Personnel, July-August 1981, p.16.

29. Doris W. Ryan and Edward S. Hickcox, Redefining Teacher Evaluation (Toronto, Ontario: OISE Press, The Ontario Institute for Studies in Education, 1980), p.vii.

potential payoffs are substantial: improved motivation, utilization and retention of personnel, and improved organizational effectiveness. The dynamic nature of the environment requires an approach to appraisal design which is iterative, proactive, anticipatory, and problem-solving; as changes in the external environment bring new data to the attention of users, the opportunity exists for greater relevance in the design and application of appraisal programs.

REFERENCES

Baird, L. S.; Beatty, R. W.; and Schneier, C. E. The Performance Appraisal Sourcebook. Amherst, Mass.: Human Resource Development Press, 1982.

Ball, Robert R. "What's the Answer to Performance Appraisal?" Personnel Administrator, July 1978, pp. 43-46.

Baroni, Barry J. "The Legal Ramifications of Appraisal Systems". Supervisory Management, January 1982, pp. 40-43.

Beach, Dale S. Personnel: The Management of People at Work. New York: Macmillan Publishing Company, Inc., 1980.

Beer, Michael. "Performance Appraisal: Dilemmas and Possibilities". Organizational Dynamics, Winter 1981, pp. 24-36.

Brinkerhoff, Derick W., and Kanter, Rosabeth Moss. "Appraising the Performance of Performance Appraisal". Sloan Management Review, Spring 1980, pp. 3-16.

Canadian Education Association. The Supervision of Teaching. Report of a CEA Seminar, Montreal, Canada, July 1979.

Cummings, L. L., and Schwab, D. P. Performance in Organizations: Determinants and Appraisals. Glenview, Illinois: Scott, Foresman and Company, 1973.

Danzig, Selig M. "What We Need to Know About Performance Appraisals". Management Review, February 1980, pp. 20-24.

Dipboye, Robert L., and de Pontbriand, Rene. "Correlates of Employee Reactions to Performance Appraisals and Appraisal Systems". Journal of Applied Psychology, April 1981, pp. 248-251.

Gruenfeld, Elaine F. Performance Appraisal: Promise and Peril. Ithaca, New York: New York State School of Industrial and Labor Relations, 1981.

Hersey, Paul, and Goldsmith, Marshall. "The Changing Role of Performance Management". Training and Development Journal, October 1980, pp. 18-22.

Kahalas, Harvey. "The Environmental Context of Performance Evaluation and Its Effect on Current Practices". Human Resource Management, Fall 1980, pp. 32-40.

Kirkpatrick, Donald L. How to Improve Performance Through Appraisal and Coaching. New York: AMACOM, 1982.

Kleinman, Lawrence S., and Durham, Richard L. "Performance Appraisal, Promotion and the Courts: A Critical Review". Personnel Psychology, April 1981, pp. 103-113.

Levinson, Harry. "Management by Whose Objectives?". Harvard Business Review, July-August 1976, pp. 125-134.

Linenberger, Patricia, and Keaveny, Timothy J. "Performance Appraisal Standards Used by the Courts". Personnel Administrator, May 1981, pp.89-94.

Locher, Alan H., and Teel, Kenneth S. "Performance Appraisal -- A Survey of Current Practices". Personnel Journal, May 1977, pp. 245-247, 254.

McMillan, John D., and Doyel, Hoyt W. "Performance Appraisal: Match the Tool to the Task". Personnel, July-August 1980, pp. 12-20.

Mintzberg, H. The Structuring of Organizations: A Synthesis of the Research. Englewood Cliffs, New Jersey: Prentice-Hall, 1979.

Morrison, Ann M., and Kranz, Mary Ellen. "The Shape of Performance Appraisal in the Coming Decade". Personnel, July-August 1981, pp. 12-22.

Rendero, Thomasine. "Consensus: Performance Appraisal Practices". Personnel, November-December 1980, pp. 4-12.

Ryan, D. W., and Hickcox, E. S. Redefining Teacher Evaluation. Toronto, Ontario: Ontario Institute for Studies in Education, 1980.

Ryval, Michael. "The Corporate Report Card". The Financial Post Magazine, March 1981, pp. 45-46.

Sashkin, Marshall. "Appraising Appraisal: Ten Lessons From Research for Practice". Organizational Dynamics, Winter 1981, pp. 37-50.

Smith, Howard P., and Brouwer, Paul J. Performance Appraisal and Human Development. Don Mills, Ontario: Addison-Wesley 1977.

Steele, James W. "In Praise of Appraisal". S.A.M. Advanced Management Journal, Summer 1981, pp. 12-19.

Teel, Kenneth S. "Performance Appraisal: Current Trends, Persistent Progress". Personnel Journal, April 1980, pp. 296-316.

Truell, George F. Performance Appraisal: Current Issues and New Directions. Buffalo, New York: PAT Publications, 1980.

Winstaneley, N. B. "Legal and Ethical Issues in Performance Appraisals". Harvard Business Review, November-December 1980, pp. 186-187.

Yankelovich, Daniel. "Yankelovich on Today's Workers: We Need New Motivational Tools". Industry Week, August 6, 1979, pp. 2-9.

4.0 Purposes of Appraisal

It has been stated that the uses of performance appraisal can be classified into two broad categories: evaluation and development. More specifically, the primary purposes of appraisal are:

* monitoring and control
* feedback and development
* compensation administration
* promotion/retention/tenure/transfers
* human resources and program planning
* organizational accountability

We will consider each of these purposes in turn.

4.1 Monitoring and control

Appraisals are an aid in maintaining a satisfactory level of performance by employees. By developing standards which articulate an organization's expectations of how an employee should perform his or her job, an employee's performance can be judged or rated against such standards in order to ensure acceptable performance and to identify deviations from expected behaviour. Where job behaviour is critically deviant from performance standards, then appraisal in some instances can be used for disciplinary purposes.

The control process in general occurs simultaneously at the individual and organizational level so that evaluation can incorporate an assessment not only of individual performance, but also of the fit or congruence of the individual's performance with the objectives, programs, and policies of the organization as a whole. Appraisal, in this way, provides the means for confirming that organizational goals or programs are being appropriately implemented by persons who are expected to contribute to their achievement.

4.2 Feedback and development

Performance appraisal is a structured way of letting subordinates know how they are doing on the job. In the behavioural sciences, the use of feedback has come to be generally recognized as an extremely important variable in employee performance. It can provide instructional information and it can motivate individuals to exert greater effort. (These are typically referred to as the directive and incentive functions of feedback.) Furthermore, management literature tells us that to be effective, feedback should be specific, frequent, direct, and constructive.

Feedback is inextricably linked with development since improvement of future performance is frequently obtained by providing each employee with periodic feedback. In addition, systematic evaluations of performance may uncover deficiencies in knowledge, skill or attitudes that can be remedied through coaching, counselling or training. Unlike the control function of appraisals, development focuses on change mechanisms to improve future performance and to identify and utilize employee potential. This is potentially beneficial to both the employee and the organization. From the individual's viewpoint, he or she is provided the stimulus for growth and the opportunity to attempt new challenges. In light of the changing social values outlined earlier, higher occupational expectations among today's work force, and workers' increasing insistence on clear knowledge of their contributions at the work place, accommodation to the benefits of individual development are particularly important. From an organizational viewpoint, individual development which is consistent with organizational needs obviously enhances organizational effectiveness.

4.3 Compensation administration

Few theoreticians would argue about the motivational potential of pay when it is tied to an employee's level of performance. Furthermore, when accurate measures of relevant behaviours are available, a clear link between pay and performance can be established. Performance appraisal provides useful data for compensation administration if proper standards or criteria of good performance can be determined and actual performance can be easily assessed. In the private sector, a recent survey of appraisal programs in a number of large American organizations revealed that "the clearest and most significant trend noted is that most organizations are making serious attempts to tie pay closely to performance."[30]

In areas such as teaching, however, where output is not easily measurable and the appropriate standards are somewhat equivocal, a pay plan grounded in merit alone is a difficult undertaking. Furthermore, formal recognition of other employee inputs (for example, education, training, experience or seniority) is often necessary.

4.4 Promotion/retention/tenure/transfers

Performance appraisals also serve as a data base for input to decisions regarding promotion, retention, tenure and transfers. The criteria utilized to make promotion decisions vary widely in practice, but tend to be based on a mixture of seniority, merit and, more covertly, interpersonal attractiveness and political behaviour. But as Glueck (1978) points out, "the more the government presses equal opportunity goals, the more likely it is that formal performance evaluation and promotion systems will be developed, validated and performance related."[31] In addition, as noted earlier, changing social values regarding employee rights are increasing the sensitivity surrounding dismissal decisions. In business for example, there has recently been a sharp increase in the number of wrongful dismissal suits coming before the courts.[32] This suggests that performance appraisal may serve a dual purpose in terms of promotion, tenure and redundancy decisions: first, as a source of information from which the organization can draw to make more objective and fair decisions; and second, to provide documentation for purposes of supporting organizational actions in the area of human resource management.

4.5 Human resources and program planning

An effective performance appraisal program permits an organization to forecast future staff needs, compare these needs with the present work force and determine the number and types of employees to recruit or let go. The appraisal system encourages supervisors to think about the abilities of their employees and to match these abilities with the known requirements of the job or jobs.[33] In terms of program planning, performance appraisal provides data on the abilities and potential of the organization's employees. The implementation of programs is facilitated by knowledge of the capabilities and deficiencies of staff resources at any given point in time. Programs requiring the use of special staff with particular strengths or abilities can be allocated to those individuals as identified by the appraisal system in order to create a compatible fit between the tasks to be performed and the individuals who will perform them.

---

[30] Kenneth S. Teel, "Performance Appraisal: Current Trends, Persistent Progress", Personnel Journal, April 1980, p.298.

[31] William F. Glueck, Personnel: A Diagnostic Approach (Dallas, Texas: Business Publications Inc., 1978), p. 323.

[32] See David Harris, Wrongful Dismissal (Toronto, Ontario: Richard De Boo Limited, 1980).

4.6 Organizational Accountability

As noted earlier, several environmental trends are converging to create pressure on public and private organizations to account to society for their actions. We have seen this trend manifested in changing legislation and social values. Performance appraisal programs provide the possibilities for joint accountability among all levels and members of an organization. Furthermore, appraisal, because it is formal and systematized, provides an objective basis upon which to support an organization's actions and to substantiate the quality of its output. Not only does the system have to be fair and equitable, it must be perceived to be so by the public. Appraisal, therefore, is essential not only as means for improving organizational performance but also as a mechanism for external accountability.

It is evident that performance appraisal can serve a number of significant uses. It is the multiple-purpose aspect of performance appraisal that is at once its greatest strength and possibly its greatest weakness. One personnel mechanism which has the capability of serving such a multiplicity of functions becomes attractive to many organizations to use in coordinating various aspects of human resource management and supervision. On the other hand, there is a small but detectable shift in attitude beginning to develop which suggests that attempting to utilize performance appraisal for a variety of functions augments its complexities and ultimately causes most of the difficulties which arise in devising an optimal system. Lazer (1980), for example, states: "the biggest reason for performance appraisal failure is too many uses and conflicting objectives."[34] Similarly, Truell (1980) claims: "some objectives are primarily oriented to past events, others are more oriented towards future events. Herein is one of the key problems found in most performance appraisal processes -- multiple objectives which are not compatible with one another."[35] Finally Yager (1981) suggests that "systems that combine all appraisal functions into one package rarely win both employee and management support."[36] What this may suggest is that initial implementation of an appraisal system should concentrate solely on one or two purposes. Its use can then be subsequently expanded in more or less modest increments to thoroughly investigate its congruence not only with the organization, but with the other purposes for which the appraisal system was initially designed.

---

[33] Valerie Stewart and Andrew Stewart, Practical Performance Appraisal (Farnborough, England: Gower Press, 1977), p.7.

5.0 Design of Appraisal Systems

5.1 Process vs. Output-Focused Systems

A fundamental consideration in the design of performance appraisal is whether the system should be process-focused or output-focused. First, let us consider what each system embodies.

A process-focused system basically concentrates on the "means to the end" of job performance, rather than the end itself. Individuals are assessed on the basis of the activities they pursue, whether they follow policies and procedures, what style they exhibit in the performance of their duties and/or the means by which they carry out day-to-day tasks.

---

34. Robert I. Lazer, "Performance Appraisal: What Does the Future Hold?", Personnel Administrator, July 1980, p. 70.

35. Truell, Performance Appraisal, p.3.

36. Ed Yager, "A Critique of Performance Appraisal Systems", Personnel Journal, February 1981, p. 130.

Boundaries are placed on the individual's degree of freedom within his or her job and performance is judged on the basis of the extent to which the individual remains within, or deviates beyond, those boundaries. Furthermore, the job descriptions from which performance standards are derived tend to be skill-based. This approach is particularly applicable in jobs where one cannot easily perceive tangible results or where evidence of output appears only after a long time span. This approach is appropriate in research and development activities and certain types of service industries for example.

In an output-focused system, as the term implies, the focus is on results. Accordingly, boundaries on the process by which results are achieved are largely discretionary or self-imposed. Job descriptions for output-focused systems tend to be results-based or achievement-based. This type of approach is appropriate for example in the case of a production manager in a manufacturing company who is given resources and a mandate to produce 'x number of widgets by fiscal year end'. The means or process by which the widgets are produced is singularly less relevant to the manager's supervisor than the manager's ability to meet production targets using a given level of resources.

The hybrid or mixed system is probably the most common in practice; that is, some minimum process boundaries are specified for any given job, coupled with a statement of expected results. The difficulty of a hybrid system is that an individual may meet the criteria on process but not those on results, or vice versa. For example, an employee may state: "I'm going by the book, I'm doing everything you tell me to do. Why can't I make the grade?" Or conversely, a manager may accomplish what was perceived as a difficult goal for his or her company but be severely criticized for having had too "authoritarian" a manner in carrying out the job. This indicates that both feedback and priorities are important considerations in the use of hybrid systems.

Problems intrinsic to process-focused systems are founded in the potential danger that the ultimate goal of the position will no longer be perceived. That is, employees may lose sight of the long-range or overall objectives of their position in the absence of measured results and either lose motivation, delay outcomes, digress from their purpose, or resent the control placed over their day-to-day performance. A common result in such systems is for employees to adhere strictly to work rules or procedures as a safety measure regardless of the consequences of such behaviour.

The problem with output-oriented systems is that they do not inform employees about whether their methods are effective or ineffective. In addition, "such performance measures are often affected by extraneous factors over which the individual has little or no control."[37] Levinson (1976) argues convincingly of the need to measure the "how" of job performance. He suggests that while systems may purport to appraise results, inevitably individuals will be judged on their methods of doing things as well. If the appraisal process focuses solely on results, judgements on the "how" of performance nevertheless will be made in a subjective and impressionistic fashion.[38] Most management literature tends to emphasize the output or results-oriented approach. Moreover, evidence indicates that " ... an emerging trend is the systematic way managers are now applying the results-oriented approach to their operations."[39] There is an obvious reason for preference of this approach: accomplishments, achievements, and results are generally easier to measure and evaluate than process factors. Furthermore, goal setting is often an inherent constituent of output-based systems and this process can be motivational to the employee.[40]

There are, in fact, both advantages and disadvantages to each system and instances where each is most appropriate. These are summarized in exhibit 8-1. Clearly, there is no "one best way" to design appraisal systems; rather, the design of a system must be tailored to the characteristics and requirements of different work situations.

---

37. Kenneth N. Wexley, "Performance Appraisal and Feedback", in Steven Kerr, ed. Organizational Behavior (Columbus, Ohio: Grid Publishing Inc., 1979), p.248.

38. Harry Levinson, "Appraisal of What Performance?", Harvard Business Review, July-August 1976, p.31.

39. Marion Haynes, "Developing an Appraisal Program", Personnel Journal, January 1978, p.16.

40. This is discussed further in section 4 on Management by Objectives.

Exhibit B-1

Appraisal System Design: Advantages and Disadvantages

SYSTEM DESIGN	ADVANTAGES	DISADVANTAGES
	Appropriate where results are not readily measurable.	Exclusive focus on process can prolong the eventual achievement of results.
Process-Focused Systems	Appropriate where tasks have a long time span for completion.	Absence of measured targets can be demotivational.
	Provides better control over implementation of programs/policies or the use of resources.	Focus on process can cause divergence from overall objectives of the job.
	Appropriate where the means to the end are irrelevant or controlled through other mechanisms.	Fails to give employee sufficient information on how to accomplish a task or correct errors.
Output-Focused Systems	Appropriate where observation of the process is impossible.	Fails to take into account extraneous factors beyond the employee's control.
	Established targets of results expected tend to be motivational.	Inappropriate if outcomes are the combined results of many employees.
	Easier to develop objective criteria of assessment.	
Hybrid Systems	Takes into account both the "means" and the "end" of any particular job.	Creates conflict of priority where individual meets process criteria but not output criteria or vice versa.

5.2 Hierarchical Differences in Appraisal Systems

In examining hierarchical differences in systems, little more need be said than that these differences exist and must be recognized in designing an appropriate appraisal system. However, this is an important point. Performance evaluation systems often malfunction because they are applied uniformly at all levels of the organization without due consideration of differences in job requirements between different levels of the organization hierarchy. It is unrealistic to expect a vice-president of an organization to be evaluated in an identical manner to an assembly worker on the shop floor. First, the standards or criteria of job performance for each type of job can be radically different. Second, the choice and number of appraisers can vary. As Glueck (1978) points out, "the higher the job in the organization, the more evaluators will be involved, however informally."[41] Third, the purposes of appraisal for different levels may vary; for example, compensation administration may only apply at higher non-unionized levels of a company. Fourth, different appraisal techniques may be applicable at different levels. Management by objectives, for example, is generally inappropriate at low levels of an organization where goal setting is beyond the control of the employee or dictated by higher authorities. Finally, the nature of the appraisal interview and the feedback generated to the employee vary considerably among levels. As mentioned previously, a "system of appraisal" rather than a single "appraisal system" should be implemented.

In education, goals of improvement in performance and professional development carry different meanings at the teacher, principal, superintendent and director levels because of variations in task and responsibility. For this reason, a system of appraisal that recognizes these differences is vital. Furthermore, variations in the size and complexity of school systems, differing needs among supervisors and teachers, and variations in instructional programs require flexibility in the design and application of performance appraisal. At the same time, however, a commonality is shared among all levels which allows the system of appraisal to be internally consistent within the organization as a whole. At all levels individuals plan, implement their plans and, it is hoped, assess the results of those plans. In addition, the ultimate purpose is to serve the needs of the student. To deliver improved educational services is the imperative which motivates all education staff. Moreover, most individuals want to work more effectively. The stimulation of growth and development is an objective acceptable at all levels.

---

[41] Glueck, Personnel, p.324.

Any system of appraisal implemented in a large organization must recognize hierarchical differences in the system. For example, output-focused appraisals are most likely to be used for managerial positions in a hierarchy and a process-focused approach used most likely at the professional service level. However, since the end result of improved performance is the same, the system can retain an internal consistency and synthesize the goals and contributions made by employees at all levels of the hierarchy.

6.0 Development of Performance Standards

6.1 Characteristics of Effective Standards

The factors by which an employee is evaluated are called the standards or criteria of the evaluation. Stated simply, performance standards establish what needs to be done and how well it must be done. The job description from which performance standards are often constructed usually describes only what needs to be done. Therefore a routine job description serves only as a departure point in developing standards. Standards are intended to translate the static job definition into criteria of dynamic work performance.

A great deal has been written on the characteristics of effective standards. Wexley (1979) has distilled these into four major parameters:

- validity: the degree to which a system actually measures what it purports to measure;
- reliability: the degree to which scores produced by any given measure are free from error;
- discriminability: the extent that performance appraisal system is able to differentiate between levels of performance of organizational members; and
- receptivity: the extent that employees perceive the performance appraisal system as accurate and meaningful.[42]

---

[42] Wexley, "Performance Appraisal", pp.241-259.

Similarly, Smith (1976) states that performance standards should be "relevant, unbiased, significant and practical".[43]

The preceding characteristics serve to describe the rudiments of effective criteria. Kirkpatrick (1982) provides us with a more specific delineation of the characteristics of effective standards; these characteristics tend to represent the opinions of most research in this area:

1. They are subject to change. Standards should be periodically evaluated and changed if certain job factors change.
2. They are based on the job and not on the person(s) in the job. Standards of performance are different from objectives. Objectives should be set for an individual rather than a job.
3. They are achievable. Practically all employees on the job should be able to reach the standard (except new employees who have not passed the probationary period).
4. They are understood by both superior and subordinate.
5. They are agreed on by both superior and subordinate.
6. They are as specific and measurable as possible.
7. They are written. Both superior and subordinate should have a copy.[44]

There is no minimum or maximum number of standards that should be set for a job. The inclusion of many standards in an appraisal program is advantageous for two reasons. First, it is generally recognized that job success is a multifaceted phenomenon. Therefore, the use of multiple criteria gives a closer approximation to the requirements of good performance. Second, if there are several criteria, the identification of specific employee strengths and weaknesses during the appraisal process is facilitated. On the other hand, the number of standards set should not be so large as to be unwieldy or impractical. In addition, if the appraisal system is being used for promotion or compensation decisions, these criteria may have to be collapsed into a single global rating for purposes of comparability among other employees or in order to tie pay increases to a single criterion.

---

[43] Patricia Smith, "Behaviors, Results and Organizational Effectiveness: The Problem of Criteria" in Marvin Dunnette, ed. Handbook of Industrial and Organizational Psychology (Chicago, Illinois: Rand McNally & Co., 1976), pp. 745-775.

[44] Donald L. Kirkpatrick, How to Improve Performance Through Appraisal and Coaching (New York: AMACOM, 1982) p.35.

Finally, because standards should be clear and agreed on, there is good reason to involve subordinates in setting standards when this is practical. The appraisee will be more motivated to meet or exceed standards if he or she perceives them as fairly established. By participation in the process, motivation is considerably more likely to occur. Moreover, the effectiveness of the standards may be enhanced by the input of more than one person to the process.

6.2 Personality, Skills and Behaviour Criteria

There are four types of standards against which to evaluate an employee's performance:

1. Personality criteria: These criteria are used in a traditional approach in which a list of traits (industrious, cooperative, intelligent, enthusiastic, conscientious, for example) are compiled, and against which an individual is assessed, usually via a graphic rating scale (discussed in section 5).
2. Skills criteria: These criteria specify particular skills and abilities which are accumulated by the individual and serve as a substitute for continuing on-the-job evidence of success (for example, a secretary who has evidenced an ability to type 100 words per minute on a typing test).
3. Behaviour criteria: These criteria relate to actions or behaviour which are expected to be exhibited if performance is satisfactory.
4. Results criteria: These criteria relate to outcomes which are assumed to be a result of the work of the employee.

6.2.1 Personality criteria

In a recent study comparing personality and "action-oriented" appraisals, Murray (1981) states that "although the trait approach has become the most widely used appraisal technique, it is also much criticized."[45] The results of the study showed that "employees appeared significantly more satisfied with the way in which they were evaluated under the objective appraisal system than under the more subjective trait approach."[46] An employee's personality is an important element in job success; yet this approach is being increasingly recognized as inferior to behavioural criteria. Clear performance standards are now basic elements necessary to comply with (U.S.) "Federal Uniform Selection Guidelines".[47]

---

[45] R. Stuart Murray, "Managerial Perceptions of Two Appraisal Systems", California Management Review, Spring 1981, p.92.

What are the problems associated with the use of personality criteria? Haynes (1978) provides a good summary:

> * The ambiguity of terms leads to appraisals which are biased by the appraiser's subjectivity and are, therefore, usually unreliable and invalid.
>
> * Although work continues in this area, it is not generally agreed which personality factors contribute, or to what degree, to an individual's performance.
>
> * Because they lack behavioural specifics, employees are generally unable to change these factors, so that including them in an appraisal system leads to antagonism and defensiveness rather than improvements.[48]

6.2.2 Skills criteria

Skill-based criteria are used much less frequently in practice than either personality or behavioural criteria. If success or failure in a job is solely dependent on the exercise of a specific skill, then this approach is feasible. Consider the work of the traditional pottery maker whose job depends only on his skill as a craftsman. Here a skill-based criterion is appropriate. But consider two other conditions which are likely to be present in our example, which minimize the possibility of inaccurate measurement. First, the craftsman has direct knowledge or feedback on the results of his work. He has only to look at his product to know if he has failed or succeeded. Secondly, there is a distinct absence of extraneous factors to influence the quality or quantity of his output. That is, his skill primarily dictates his output and because he works autonomously with few tools, he has considerable control over his work environment. In the presence of these conditions, more sophisticated performance measures become unnecessary.

---

[46] Ibid., p.93.

[47] See, for example, Marshall Sashkin, "Appraising Appraisal: Ten Lessons From Research to Practice", Organizational Dynamics, Winter 1981, p.44.

[48] Haynes, "Developing an Appraisal Program", January 1978, p. 16.

Skill-based criteria can also be appropriately used for appraisal of persons in entry-level or probationary jobs. Normally, the possession of skills by a professional is demonstrated by credentials or selection tests prior to hiring, but how a person actually uses those skills in their work may have to be assessed during a probationary employment period. Similarly, a person may be well qualified and knowledgeable in certain subject areas, but his or her teaching skills may have to be developed through on-the-job training or teaching experience. Before a decision is made to offer a person a more permanent status with the organization, an assessment has to be made of their ability to use or ability to learn an array of necessary skills. For this sort of assessment, a skills-based appraisal process is appropriate.

In a number of innovative approaches to work design,[49] pay increases and promotion have been linked solely to the acquisition of certain skills in a job. As each skill is mastered, the employee becomes eligible for a larger salary and added responsibilities. Although this approach has been used in certain industrial situations, factors beyond the employee's control tend to confound the individual's ability to acquire new skills. Also acquisition of a skill does not ensure an acceptable change in quality or quantity of output. Finally, there is also the problem of limited upward job mobility and pay raises once the worker has mastered all skills available in his or her work environment. For these reasons, and because certain conditions must be present for skill-based criteria to be relevant, this approach has not been extensively applied in practice in industry. Pay increments for acquisition of credentials is fairly common in education systems, but these increments are rarely contingent on changes in performance.

---

[49] See, for example, Richard E. Walton, "Work Innovations in the United States", Harvard Business Review, July-August 1979, pp. 88-98.

6.2.3 Behaviour criteria

Behaviour criteria are similar to skills-criteria but possess greater potential for meeting the four characteristics of effective standards outlined earlier: validity, reliability, discriminability, and receptivity. However, research on and application of type of criterion are relatively recent in comparison with the length of time appraisal programs have been in practice. The advantage of behavioural criteria is that they attempt to overcome the weaknesses inherent in personality or trait criteria outlined above. The development of behaviourally-based standards is considerably more difficult and time-consuming than the trait approach, but the pay-offs, in terms of greater objectivity and validity, appear to outweigh the added effort required. Behaviourally Anchored Rating Scales (BARS), one of the techniques discussed in the following section, will serve to elucidate some of the strengths and weaknesses of behavioural standards. These and other techniques will be discussed later in section 7.0.

6.2.4 Results criteria

Results-criteria for performance appraisal have the potential for satisfying many of Kirkpatrick's (1982) characteristics of effective standards which were described previously. Results-criteria can be based on the job, be specific and measurable, be time-oriented, and can be written and subject to revision. Kirkpatrick also requires that effective standards of appraisal be achievable, understood and agreed upon by both superiors and subordinates. It is this latter aspect which causes problems with many results-criteria systems because supervisors tend to impose unilaterally a set of result requirements which may be neither achievable, understandable nor agreed upon. Thus there are two important aspects in using results-criteria for appraisal: the process by which such criteria are developed and the specificity of criteria resulting from the process.

Typically, results-criteria are used in output-focused appraisal systems and as discussed previously, such systems are common in organizations which produce tangible, countable or measurable outputs. The appropriateness of output-focused appraisals in professional service organizations has been a subject of controversy due to the difficulty of developing results-criteria for this type of organization. The key to the controversy is the problem of establishing a direct causal link between the work of the professional and the effects of the service on the client. It has been effectively argued by opponents of output-focused systems that too many variables interfere with the causal relationship between professional service work and client effects for an output-focused system to be valid. Second, it has been argued that results-criteria are difficult to develop which are valid in actually reflecting the true quality or volume of professional service. Relevance of results-criteria is also a subject of debate. For example, consider the cliché "the operation was a success, but the patient died". The first part of the statement refers to results-criteria which are concerned with the technical aspects of the process of professional service work. The second part of the statement refers to results-criteria which emphasize a different aspect of performance. Which set of results-criteria is relevant will depend upon either the goals or value systems of the organization.

Finally, there is the well-known phenomenon of goal displacement in organizations as described by Etzioni: people work to satisfy results-criteria rather than to achieve the goals of their organization or profession.[50] This problem can occur when performance is appraised exclusively on ability to satisfy results-criteria, and in self-defense, people shift their priorities to look good on the criteria. Goal displacement occurs when the results-criteria turn out to be a poor reflection of or have low validity to the goals of the organization or profession.

Despite these inherent difficulties in the use of results-criteria appraisal systems, their potential to satisfy standards such as Kirkpatrick's makes this approach attractive. If an organization is able to develop such criteria so that they provide a valid indication of some aspects of a professional's performance level, then a great deal will have been achieved. On this last point, it should be recognized that the various types of criteria discussed in this section are not mutually exclusive and that different sorts of criteria can be used for appraisal of different aspects of a professional's work.

---

50. Amitai Etzioni, Modern Organizations (Englewood Cliffs, New Jersey: Prentice-Hall, 1964).

7.0 Appraisal Techniques

7.1 Classification of Techniques

Performance appraisal techniques can be differentiated in several ways. One distinction is by what is measured; that is, whether the technique emphasizes process or results. This was discussed in section 3.1. A second way to distinguish between performance appraisal techniques is by the uses to which the appraisal results are put. The particular purpose of appraisal -- judgemental or developmental -- will impact upon the techniques chosen and the manner in which appraisals are implemented. Since the two approaches can be in conflict, it may be that two or more methods will have to be utilized or combined in order to optimize the appraisal system.[51]

A third method of differentiating between appraisal techniques (which is the method of classification adopted in this study), is to consider whether the individual's performance is measured individually against some absolute standard, or in a group, compared to the performance of others. Individual or absolute appraisal techniques include the following: essay evaluation, graphic rating scales, behaviour checklists, forced choice techniques, and critical incident techniques. Multiple person or relative techniques discussed are: ranking, paired comparisons, and forced distributions. In addition, certain more sophisticated and relatively recent techniques are examined including: behaviourally anchored rating scales (BARS), management by objectives (MBO), and assessment centres.

Many of the advantages and weaknesses of the different techniques are outlined in the remainder of this section. The choice of which technique to use will depend on a variety of factors, including the size and complexity of the organization, the purposes to be served by the appraisals, and the relative costs involved. Exhibit B-2 reproduces a table of criteria for the choice of performance evaluation techniques, developed by Glueck (1978).[52]

---

51. An example of a "systems approach" to appraisal is that used by Corning Glass Works, which combines MBO results for judgemental decisions with critical incidents performance review for developmental decisions. See Michael Beer and Robert A. Ruh, "Employee Growth Through Performance Management", Harvard Business Review, July-August 1976, pp.59-66.

52. See Glueck, Personnel, p.134.

7.2 Individual or Absolute Appraisal Techniques

7.2.1 Essay evaluations

In this technique, the evaluator is asked to write a narrative of the employee's strengths, weaknesses and future potential. While the evaluator can be provided with guidelines, the essay approach tends to be fairly unstructured and difficult to use. Apart from the relatively low costs of development, the major advantage of the essay technique is the rich source of feedback which it can provide to the employee. However, since essays can vary widely and cover different aspects of performance, they are of little use in making meaningful comparisons for the purpose of judgemental decisions. For such a purpose, there is a need for more objective rating or ranking. Thus, in many organizations essays are used for developmental purposes to supplement some other method, such as a graphic rating scale.

Exhibit B-2

Criteria for Choice of Performance Evaluation Techniques

Evaluation Base	*Graphic rating scale*	*Forced choice*	*MBO*	*Essay*	*Critical incidents*	*Weighted checklist*	*BARS*	*Ranking*	*Paired comparison*	*Forced distribution*	*Performance test*	*Field review*
Developmental cost	Moderate	High	Moderate	Low	Moderate	Moderate	High	Low	Low	Low	High	Moderate
Usage costs	Low	Low	High	High supervisory costs	High	Low	Low	Low	Low	Low	High	High
Ease of use by evaluators	Easy	Moderately difficult	Moderate	Difficult	Difficult	Easy	Easy	Easy	Easy	Easy	Moderately difficult	Easy
Ease of understanding by those evaluated	Easy	Difficult	Moderate	Easy	Easy	Easy	Moderate	Easy	Easy	Easy	Easy	Easy
Useful in promotion decisions	Yes	Yes	Yes	Not easily	Yes	Moderate	Yes	Yes	Yes	Yes	Yes	Yes
Useful in compensation and reward decisions	Yes	Moderate	Yes	Not easily	Yes	Moderate	Yes	Not easily	Not easily	Yes	Yes	Yes
Useful in counseling and development of employees	Moderate	Moderate	Yes	Yes	Yes	Moderate	Yes	No	No	No	Moderately	Yes

Reproduced from: William F. Glueck, Personnel: A Diagnostic Approach (Dallas, Texas: Business Publications, 1978), p.134.

7.2.2 Graphic rating scales

One of the oldest and most widely used appraisal techniques is the graphic rating scale. This method involves rating various characteristics of the employee's performance along some graphic scale or continuum. The scale can range in specificity from a simple "unsatisfactory-satisfactory-outstanding" continuum to more specific descriptions of different levels of performance. Scales can also vary in the number of characteristics evaluated and in the degree to which scale points are quantitatively anchored for purposes of factoring or reaching an overall assessment of the employee. In order to provide a well-structured appraisal, there must be a clear definition of the performance dimensions used as standards in order that the scale become a meaningful, unambiguous and consistent tool for both the appraiser and the individual interpreting the results.

Although graphic rating scales do not yield the depth of a narrative essay and are more expensive initially to develop, they do offer several advantages. First, the administration of the scale technique is less time-consuming and costly than other techniques, such as essays or critical incidents. Second, when structured properly, the results are easy to understand and interpret. Finally, results can be quantified and standardized, thus permitting comparison between individuals and groups. However, despite these advantages, graphic scales suffer from a number of weaknesses, the most prominent being the high potential for rater error.

As with any appraisal technique that gives a high degree of control or leeway to the evaluator, the graphic rating scale is vulnerable to common rating errors such as:

* leniency: the tendency to give predominantly higher ratings, particularly where negative ratings must be explained to the subordinate;
* central tendency: the tendency to rate all employees within a narrow range;
* halo effect: the tendency to form an overall assessment of a person and then assign similar ratings to each factor, consistent with the overall assessment;
* recency effect: the tendency to remember, and therefore give more weight to, the most recent behaviour of the appraisee; and
* bias: the tendency to colour ratings through individual rater biases and subjectivity.

Although these errors may be somewhat controlled through proper instruction of the evaluators, they constitute the most widely criticized weakness of this technique. Nevertheless, rating scales remain the most commonly used appraisal technique and, particularly when supplemented by meaningful narrative comments, have proven to be as consistent and reliable as many more elaborate techniques.

A variation of the graphic rating scale is the mixed standard scale developed by Blanz and Ghiselli (1972).[53] In this technique, the rater is given three statements to describe a particular performance characteristic representing, respectively, good, average, and poor performance. By measuring each employee against these standards, the method yields a seven-point scale which Blanz and Ghiselli contend provides a more valid rating. In particular, the mixed standard scale has been shown to reduce halo and leniency errors, as well as central tendency.[54] Another advantage is that the scale can be used to evaluate the effectiveness and consistency of different raters. The primary weakness of this method is that it is time consuming to prepare and administer and accordingly, is likely to be resisted by the evaluator. In addition, the nature of the "standard" statements is such that they may not provide meaningful feedback for developmental purposes.

7.2.3 Behaviour checklists

In this method, the evaluator is provided with a list of descriptive statements. He or she is asked to check those which are applicable to a particular employee. In a variation of this method, each statement can be scaled so that the employee is rated on the extent to which he or she fits or complies with the description. In this case, each individual rating can be totalled for an overall score in accordance with the dimensions of the scale.

In order to prepare a well-constructed scale, initial development can be time-consuming and costly, but once prepared, actual administration is relatively easy. Common weaknesses include arbitrary choice of response categories or descriptive statements, unequal intervals between scale points, lack of regard to the value judgements inherent in each statement, ambiguous statements or scale points, and too many statements or scale points.

---

53. F. Blanz and E. E. Ghiselli, "The Mixed Standard Scale: A New Rating System", Personnel Psychology, 25 (1972), pp.185-199.

54. Ibid.

In its more complex form, the weighted checklist becomes similar to the graphic rating scale. Since it yields an overall summary rating, it is less vulnerable to the halo effect. However, there remains little control over other rating errors, such as leniency. The other disadvantage is that like the graphic rating scale, the results do not provide meaningful feedback for development purposes.

7.2.4 Forced-choice

The forced-choice method was designed to counteract the tendency toward higher ratings (leniency) among evaluators. In this method, descriptive statements are grouped together in such a way as to appear equally attractive or unattractive to the rater. The rater does not know which of the statements reflects good or bad performance relevant to the characteristics of the position. For example, from the following two statements,

(a) Patient with slow learners
(b) Lectures with confidence,[55]

the rater must determine which best describes a subordinate whose job is that of an "instructor". In this example, it was found by research that only (a) was relevant to effective job performance but since the rater does not know this, the tendency toward leniency is avoided. This is a primary advantage of the forced-choice method. However, since the rater has no control with this system, there can be some resistance to this technique.

As with rating scales and weighted checklists, the development costs can be high, since statements must be derived which appear equally desirable but which effectively discriminate between acceptable and unacceptable performance. Once designed and tested, however, the system should provide a valid means of differentiating between individual levels of performance. The forced choice technique is of little use for developmental purposes though, since the results are difficult to understand both for the rater and ratee.

---

[55] Example derived from Baird, Beatty and Schneier, eds. The Performance Appraisal Sourcebook, p.45.

7.2.5 Critical incidents

In this method, critical job behaviours are described in the context of specific incidents observed by supervisors and/or personnel specialists. These critical incidents, then, define behaviours which differentiate between effective and ineffective performance in a particular job. Once recorded and defined, the list of incidents provides the basis for evaluating an employee on all critical aspects of the job. Using the categories developed and criticial incidents identified, the evaluator can keep a log of each employee's behaviour, noting the frequency of the employee's critical incidents in each performance category.

The technique avoids much of the common rater errors referred to above, since the rater can focus on specific incidents and actual job behaviours rather than on general descriptions of traits or expected performance. The critical incident method also permits a basis for meaningful feedback because the ratee can visualize specific incidents and areas where change or improvement is necessary. In addition, because the supervisor keeps a running log of each employee's behaviour, there is little chance of recency error.

The weaknesses of this approach include the high costs of development, documentation, and administration. Since the technique involves maintaining a running log, the supervisor must be willing to commit time to the appraisal process on an on-going basis. If observations are less frequent, then feedback for developmental purposes is jeopardized. Moreover, if the critical incidents are formulated without any input from the employees, there may be less likelihood of acceptance by those evaluated. Finally, the results tend to be qualitative in nature and therefore are less useful than quantitative measures in making judgemental comparisons between different employees and employee groups.[56]

---

[56] A modification of this technique, designed to overcome some of these problems, is the behaviourally anchored rating scale (BARS), discussed later in this section.

7.3 Multiple Person or Relative Appraisal Techniques

7.3.1 Ranking

In the simplest form of this method, the rater is expected to rate employees from highest to lowest on some criterion or criteria of effectiveness. A variation of this is alternative ranking, in which the rater first chooses the best and worst employees, then the second best and second worst, and so forth, moving toward the middle until all employees have been ranked. The greatest advantage of these techniques is their simplicity; that is, they are easy and inexpensive to develop, administer and understand. They are useful in personnel decisions involving direct comparisons among employees, but do not provide meaningful feedback for developmental purposes. A final advantage is that, because employees are ranked from highest to lowest, there are no central tendency or leniency errors.

Despite their simplicity, there are a number of weaknesses inherent in ranking techniques. First, ranking does not provide information as to the degree of difference between employees. Second, a ranking on the basis of one global criterion may be invalid because it lacks a sound behavioural base. Third, the ranking method can lead to unfair treatment of employees who receive a low ranking in a superior group, but who may be more effective than those with a high ranking in an inferior group. Similarly, the ranking method does not permit meaningful comparisons across different groups or departments. Finally, there is a tendency for rankings at the extreme ends of the scale to be more reliable than those in the middle. Rankings made by different raters and even by the same rater may vary at different times.

7.3.2. Paired comparisons

Some of the reliability problems intrinsic to ranking methods can be overcome by use of the paired comparison technique. In this approach, all employees to be evaluated are listed on a card so that each individual is compared directly against each other individual. In each case, the evaluator must simply choose the more effective of the two employees in accordance with some overall criterion. The end result is a more reliable ranking, based on the number of times a particular employee is preferred over the others. Provided the number of appraisees remains relatively small, this method remains easy and inexpensive to develop, administer, and understand. However, apart from the improved reliability, paired comparisons are subject to the various weaknesses identified earlier with respect to ranking.

7.3.3 Forced distribution

Another comparison method, designed to reduce leniency and central tendency errors, is that of forced distribution, in which the appraiser must categorize employees in terms of some fixed distribution, such as 10% superior or 20% good, for instance. As with the ranking and paired comparison methods, this approach can also result in unfair assessments. Consider the low 10% in a superior group, for example. In general, forced distribution would not be more advantageous than the other comparison methods except, perhaps, in respect of compensation decisions in which it is relevant to categorize employees in different group levels rather than ranking each employee in an ordinal manner.

A more recent variation of forced distribution is the point allocation technique in which the evaluator is given a quota of points per employee and the total number of points allocated cannot exceed the number of points per employee times the number of employees evaluated. This method gives some indication of relative distance between employees, but remains subject to many of the other criticisms outlined above.

7.4 Other Appraisal Techniques

7.4.1 Behaviourally Anchored Rating Scales (BARS)

This technique is a modification of the critical incidents approach which is designed to produce valid performance criteria while minimizing both appraiser error and resistance.

Preparation of this appraisal instrument is significantly more sophisticated and time consuming than other techniques. It essentially involves five major steps. First a group of workers and/or supervisors is asked to describe critical incidents of effective job behaviour. Second, these incidents are clustered into smaller sets which define important job dimensions. Third, for verification purposes, a second group of people is given the same incidents and performance dimensions and is asked to reclassify each incident into the appropriate dimension. Only those consistent with the first group's classification are retained. Fourth, the second group rates each incident according to the degree to which the behaviour described represents effective or ineffective performance based on the relevant job dimension. Ratings are retained if there is significant consistency among the raters. Finally, the incidents which have been reclassified and retained are used as behavioural anchors on the rating scale for each job dimension. Thus the final instrument consists of a rating scale, anchored by the retained critical incidents.[57]

Because the development of BARS instruments is costly and must be repeated for different jobs, this approach is generally impractical for many organizations. Nevertheless, it does promise several advantages over other appraisal techniques. First, the scales have a sound basis in actual job behaviour rather than generally described performance characteristics or value traits. The process by which the instrument is designed ensures that the performance dimensions are well-defined and that the criteria are job relevant and hence, valid. Secondly, the high degree of participation by both subordinates and supervisors should contribute to acceptance and commitment at both levels. Third, the scales provide a meaningful source of feedback for developmental purposes. Finally, the method promises to reduce rater errors such as halo effect and leniency, through employee participation and the use of critical incidents to anchor scale points in behavioural terms.

Unfortunately, despite the numerous apparent advantages, the use of BARS has not become widespead and the research to date has not shown it to be significantly better than other rating formats. In particular, most studies have indicated that BARS do not yield a significant improvement over other rating techniques in such areas as rater error, inter-rater reliability and validity.[58]

---

[57] For a more complete description of BARS development, see Herbert G. Heneman and Donald P. Schwab, Perspectives on Personnel/Human Resource Management (Homewood, Illinois: Richard D. Irwin, 1978), pp.64-66.

[58] See, for example, Paul O. Kingstrom and Alan R. Bass, "A Critical Analysis of Studies Comparing Behaviourally Anchored Rating Scales (BARS) and Other Rating Formats", Personnel Psychology, Summer 1981, pp.263-289.

7.4.2 Management by objectives

Management by objectives (MBO) diverges radically from other forms of appraisal in that, rather than comparing actual job performance to some predetermined standards or criteria of excellence, it is measured against goals prescribed jointly by the appraiser and the appraisee. In essence MBO requires that an individual and his or her superior determine a precise statement of performance objectives to be accomplished at a particular date by the subordinate. Appraisal is then based on the degree to which the individual is able to achieve these goals. Performance is compared against these objectives at the stated accomplishment date and achievements, as well as variances, are discussed. Subsequently, new goals are jointly established and the cycle is repeated. MBO then possesses three main elements, namely, goal setting, superior-subordinate participation, and feedback.

A great deal of literature has surfaced in recent years advocating or criticizing the use of MBO as an appraisal technique. Proponents suggest that a number of advantages accrue to organizations utilizing an MBO program. First, there is an increase in the quality and quantity of communication between superiors and subordinates. Second, goal-setting and feedback are motivational to the employee and hence, improve performance levels. In addition, MBO permits individuals greater freedom and control in managing their own work. This is not only satisfactory to the employee, but allows the individual to be more proactive in affecting and improving his or her work environment. The result is a more satisfied and motivated employee. Third, MBO is useful as a planning tool at both the organizational and individual level because it focuses on future objectives and brings about an identification of problem areas in goal attainment. Finally, with the emphasis on results, there is a reduction in appraiser error and subjectivity with regard to evaluating subordinate performance.

An almost universally cited criticism of MBO is the substantial amount of time required to practice it. In addition, superiors often resist MBO not only because it may be an administrative burden, but also because it may, through the process of joint decision-making, represent an erosion of their authority. Second, MBO may serve in some instances to discourage initiative. Individuals may refuse to set difficult goals, believing that the successful attainment of moderately difficult goals will be perceived more favourably than failure at more difficult objectives. Third, while the one-to-one relationship between superior and subordinate is fostered under MBO, this may serve to weaken the links between the employee and other members of the organization; that is, the emphasis on specific individual goals may dilute the need for organization interdependence among all staff members. Fourth, similarly, MBO can create problems in the coordination and integration of efforts. Self-established goals and job tasks may turn out to be overlapping or basic tasks may be unperformed. Furthermore, individual goals may become incongruent with organizational objectives if the goal-setting process is not carefully controlled and monitored. Finally, and perhaps most importantly, "the measurement process tends to partially distort the realities of the job."[59] This appears to be the case for two reasons. First, "there is a tendency to establish objectives in areas where measurement of accomplishment is relatively easy, especially quantitative measurement. These may not be the most important areas."[60] Secondly, as discussed in section 5.1, it is not always feasible to focus specifically on results or outcomes, with minimal attention paid to the process by which those results are achieved. Relevant behavioural characteristics can be overlooked, employees may act in ways that are not in keeping with the organization's policies, style, procedures or climate, certain extraneous factors beyond the employee's control may influence the attainment of objectives, and the employee may receive inadequate information on the process by which he or she is expected to reach established goals.

In considering MBO as a possible appraisal technique, it should be kept in mind that an MBO program is often part of a more systematic program of organizational change, and usually involves all levels of the organization. Implementation of MBO is a complex and lengthy process that requires full organization support through proper training of personnel, linkages between the MBO program and the reward system, top-down support at all supervisory levels, and continual, systematic monitoring of the program. Personnel at all levels of the hierarchy must be fully committed to MBO in order for this technique to be effective.

7.4.3 Assessment centres

The key concept in assessment centres is the simulation of critical job tasks in a special environment in order to test a candidate's capabilities to perform in the job. Job analysis is used to identify the critical job tasks and then a series of simulations of these tasks are created. Candidates perform the simulated tasks for an intensive period of time while several centres can be very expensive to develop and maintain (for reliability and validity of criteria), but can be very useful for jobs where probationary trials or learning-on-the-job is risky or costly.

---

[59] John B. Miner and Mary Green Miner, Personnel and Industrial Relations (New York: Macmillan Publishing Company, Inc., 1977), p.194.

[60] Ibid.

The use of assessment centres in industry is generally limited to quite large employers such as AT&T, General Electric, Alcan and Northern Telecom. The basic purpose of assessment centres is similar in most organizations: the identification of skills and abilities or the assessment of candidates' potential for different types of work. Jaffee and Cohen (1978) have outlined the following major differences between assessment centres and performance appraisal:

* Assessment centres focus on anticipated job requirements while performance appraisal focuses on current job requirements.

* Assessment centres use a simulated work environment with concurrent observation of performance while performance appraisal uses the actual work situation with historical reviews of performance.

* Assessment centres use multiple raters looking at specific skills while performance appraisal generally uses a single rater looking at general performance.[61]

In summary, assessment centres are most appropriately used for selection and possibly for assessment of training needs, but are not normally used for performance appraisal.

8.0 Implementation of the System

8.1 Choice of Appraiser

The first question that one poses in implementing an appraisal system is: Who should conduct the appraisal? The current practice in most organizations is for individuals to be evaluated by their supervisors. There are, however, five sources of appraisal:

---

[61] C. L. Jaffee and S. L. Cohen, "Improving Human Resource Effectiveness Through Assessment Centre Technology", in E. L. Miller, ed., Management of Human Resources (Englewood Cliffs, N.J.: Prentice-Hall, 1980), pp. 350-379.

1. supervisory appraisal
2. peer appraisal
3. appraisal by subordinates
4. self appraisal
5. appraisal by outsiders

We shall consider each of these in turn.

8.1.1 Supervisory appraisal

There are three primary justifications for appraisal by the supervisor. The authority structure intrinsic to most organizations legitimates the control of the supervisor over the subordinate. It is therefore reasonable that, for both evaluative and developmental purposes, responsibility for the assessment of performance should reside with the supervisor. Second, the supervisor is often in a position to determine the magnitude or frequency of rewards and punishments allocated to various subordinates. To the extent that rewards can be tied to the performance level of subordinates, both the assessment of performance and the distribution of rewards should be allocated to the same decision-making authority. Third, it is often assumed (although this is not always the case) that the supervisor has the greatest opportunity to observe the behaviour of the appraisee.

What then are the potential liabilities of supervisory appraisal? The major problem is the appraisal communication process. Subordinates may feel threatened, defensive or reluctant to communicate openly with a higher authority. In addition, as Cummings and Schwab (1973) point out:

> The superior frequently feels uncomfortable in the appraisal role because: (1) the role demands skills which he does not possess, (2) he has an ethical objection to "playing God", (3) communication of negative appraisals may alienate subordinates, or (4) he realizes that in some cases the superior does not have sufficient reward-and-punishment power to implement the results of an appraisal even if he does a thorough job of it.[62]

These potential problems are often manifested in subordinates' resistance to appraisal and superiors' avoidance of or delay in conducting the appraisal.

---

[62] L. L. Cummings and Donald P. Schwab, Performance in Organizations: Determinants and Appraisal (Glenview, Illinois: Scott, Foresman and Company, 1973), p.104.

All these problems are not necessarily irreconcilable. If an appraiser's own performance is partially evaluated and rewarded on the basis of his or her ability to assess his or her subordinates, then the appraiser will be motivated to conduct better appraisals. In addition, some skill deficiencies can be rectified by appropriate training. Furthermore, the organization can allocate some reward-and-punishment power to the supervisor to aid in implementing appraisal results. Finally, in instances where the supervisor has limited opportunity to observe the subordinates' behaviour, other sources of appraisal (self-appraisal, peer appraisal) can serve to supplement the supervisor's evaluation.

8.1.2 Peer appraisal

Peer appraisal can be effective if the evaluating peers trust one another and are not competitive for raises or promotions. This approach is most appropriate in situations where there is a great deal of opportunity to observe the performance of colleagues. Lewin and Zwany (1976) have reviewed the research on peer ratings and found many more positive than negative studies for peer evaluation. Yet peer evaluations are rarely used in practice. The reasons for this may be associated, as mentioned above, with the authority structure which exists in most organizations. Peer ratings simply run counter to notions of supervisory responsibility and accountability ability for the performance of subordinates. Moreover, there may only be limited areas in practice where peers do not compete for similar organizational rewards or use of resources. Finally, peer group pressures often create certain biases which cause peers to give distorted or inaccurate appraisals.

8.1.3 Appraisal by subordinates

This approach to appraisal is used very infrequently except in university settings, for example, where students may evaluate faculty. In addition, it is used more for the developmental aspects of appraisal than some other methods. As Wexley states:

> Appraisal by a manager's subordinates is an excellent developmental technique. It provides managers with meaningful feedback about how their people perceive them and what behaviors they should modify or improve. However, subordinate appraisals should rarely be used for salary or promotional decisions. Many managers consider subordinate appraisals to be highly unorthodox, undermining to their legitimate power, and contaminated by subordinate self-interest.[63]

---

63. **Wexley, "Performance Appraisal", p.243.**

Subordinate appraisals would appear to be appropriate if: the supervisor cannot retaliate against a subordinate who gives a poor rating (through mechanisms of anonymity, for example); the subordinate's evaluation cannot or will not be coloured or biased by desires for aggrandizement; it is the intent of the organization to create more power equalization between superiors and subordinates; knowledge of perceptions of subordinates will provide useful data for developmental purposes and motivate the supervisor to change his or her behaviour; or information either cannot be obtained from other sources or is insufficient by itself to provide a complete and realistic assessment of performance.

8.1.4 Self-appraisal

Self-appraisals appear justified in circumstances where the individual must work in isolation, where it is difficult or impossible for others to observe the performance of the individual, where the individual is the unique possessor of a rare skill, or where there is a strong emphasis on development, personal growth, and self-motivation of the individual.

Several positive results have been found to be associated with self-appraisal, "including: (a) more satisfying and constructive appraisal interviews, (b) less defensiveness by performers regarding the appraisal interviews as well as the overall appraisal process, (c) and improved job performance."[64]

There are, however, problems associated with self-appraisal. Research has generally shown a low agreement between self- and supervisory appraisals, and a tendency for self-appraisals to be inflated, particularly among poorer performers.[65] It may be that what is evaluated and why will influence supervisory-subordinate agreement. If standards are behaviourally-based rather than trait-based, there is less latitude for disagreement stemming from ambiguous criteria. Moreover, if the purpose of appraisal is developmental and non-evaluative, it is less likely that self-appraisals will be subject to systematic biases and distortions.

---

[64] Cummings and Schwab, Performance in Organizations, p.106.

[65] See, for example: H. G. Heneman, "Comparisons of Self- and Superior Ratings of Managerial Performance", Journal of Applied Psychology, 59 (1974), pp.638-642; or R. J. Klimoski and M. London, "The Role of the Rater in Performance Appraisal", Journal of Applied Psychology, 59 (1974), pp.445-451.

It is evident that self-appraisals are effective tools for encouraging self-growth and development. It seems inevitable that an individual will periodically undergo some form of self-appraisal, if only for himself or herself, and especially during an appraisal interview with a supervisor. Therefore, where it is appropriate, it may be wise to make the process of self-appraisal more formal or systematic, in order to assist the individual in thinking about his or her own performance in meaningful and constructive ways.

8.1.5 Appraisal by outsiders

Appraisal can be conducted by personnel specialists, trained resource people within the organization, or external consultants (psychologists, for example). This approach may be appropriate if someone with specialized expertise, training or knowledge in a particular content area has the capability of making an accurate assessment. Alternatively, an external appraiser may be required in situations where the vested interests of other sources are too extreme to permit a fair and reliable appraisal. Here the objectivity of the assessment can only be ensured by calling in a third party.

Advocates of appraisal by outsiders argue that this method increases reliability and reduces variances in appraisal by the specialist applying a common frame of reference across similar jobs and situations. In addition, it is sometimes claimed that use of an external person saves the supervisor time and permits him or her to avoid what may be perceived as a distasteful chore. There are nevertheless two major disadvantages to this approach:

> First, it encourages the superior to avoid an essential part of his managerial reponsibility. Second, it is questionable whether meaningful conversation will take place between a superior and his subordinates about performance improvement if the evaluation is conducted by an outside party.[66]

---

66. Cummings and Schwab, Performance in Organizations, p.109.

It is evident that the choice of appraiser must be contingent upon the nature of the job and the work environment. The advantages and disadvantages inherent in each approach must be weighed carefully to determine the optimal source of appraisal. It may be that no one organizational source can provide all relevant information for assessing employee effectiveness: there is growing evidence that multiple appraisers can increase the validity and reliability of the appraisal process. For many organizations, then, the best approach may be a combination of two or more appraisal sources.

8.2 Appraiser Training

In a recent review covering the past 25 years of research literature on appraiser training. Spool (1978) concluded:

> The state of the art in training observers of behavior appears to be in its infant state. To date not much research has been conducted and of the research that does exist, most have serious methodological flaws, and extremely few are comparative or systematic in nature. As a result, very little is known about which training approach is most effective in increasing accuracy of observation and which components of the training design contribute most to the overall effectiveness of the training program.[67]

The main purpose of appraiser training is to increase the objectivity and accuracy of the evaluator's assessment by reducing such errors as halo effect, leniency, central tendency and bias described earlier. Rating errors are difficult to correct because observers are usually unaware that they are making them.

Recent research has concluded that if rater training programs are to be effective they should emphasize:

* the multidimensionality of work performance
* the importance of recording objectively what is seen
* the development of specific examples of effective and ineffective performance.[68]

While it is not within the scope of this discussion to outline in detail the content of various training programs, it is important at this point to draw attention to two sources which may hold particular promise in guiding future training programs for appraisal purposes.

---

67. Mark D. Spool, "Training Programs For Observers of Behavior: A Review", Personnel Psychology, 31 (1978), p.888.

68. See H. J. Bernardin and E. G. Pence, "The Effects of Rater Training: Creating New Response Sets and Decreasing Accuracy", Journal of Applied Psychology, 65 (1980), pp.60-66.

Bernardin and Buckley (1981) have developed a framework for rater training programs in which three areas are particularly emphasized:

(a) diary-keeping procedures to increase observational skills;

(b) the establishment of a common rater frame of reference to enhance agreement on what constitutes effective job performance; and

(c) mastery-based training to increase rater self-efficacy regarding negative situations.[69]

The interested reader is referred to this source for further elaboration. In addition, Latham, Wexley and Pursell (1975) have developed a performance appraisal training program[70] which has been described as "the most advanced rater training program related to rating job performance".[71] This training occurs in a workshop in which participants see videotapes of hypothetical candidates being appraised by a supervisor. The trainees are asked to rate each candidate and to determine how they think the supervisor in the videotape rates each candidate. By this process, trainees have the opportunity to observe rater errors (on the videotape), actively participate in learning to rate accurately themselves, receive feedback regarding their own rating abilities, and practice to reduce any errors they may commit.

Historically, the subject of evaluator training has centred upon the attempted reduction of rater error and rater bias in the actual process of evaluation. More recently, thought is being increasingly given to the expansion of evaluator training programs to encompass interpersonal skills and methods of feedback. There is a growing recognition that evaluator training should instruct the appraiser in how to handle appraisal interviews effectively, how to give constructive feedback and how, when appraisal is used for both developmental and evaluative purposes, to handle the dual role of "judge" and "helper" successfully. At the same time that this need for training in interpersonal relationships is being recognized, little is presently being done to translate need into valid and workable training modules. Furthermore, there are those who suggest that the success of such training would be limited because these types of skills are not easily learned.

---

69. H. John Bernardin and M. Ronald Buckley, "Strategies in Rater Training", Academy of Management Review, April 1981, pp. 205-211.

70. G. P. Latham, K. N. Wexley and E. D. Pursell, "Training Managers to Minimize Rating Errors in the Observation of Behavior", Journal of Applied Psychology, 60 (1975), pp.550-555.

71. W. C. Borman (1979), cited in Baird, Beatty and Schneier, The Performance Appraisal Sourcebook, p. 88.

Research evidence indicates that appraiser training is still "in its infant stage". Yet recent studies, like the two mentioned above, suggest that valuable efforts are being initiated to increase the accuracy of evaluation and reduce rater errors. Steele (1981) claims that " ... too many programs have failed because sufficient user training has not taken place. This is perhaps the single most important factor influencing the failure of most performance appraisal programs."[72] If this is, in fact, the case, future directions in performance appraisal will have to increasingly emphasize this type of support for the appraisal process. However, as the design and techniques of appraisal systems become more sophisticated, more objective and more independent of rater bias, the imputation of fault to the evaluator for causes of appraisal failure will diminish.

8.3 Assessment Period and Timing

This aspect of performance appraisal systems considers the frequency of appraisals and the optimum time interval between appraisals. There are basically four approaches outlined which organizations can utilize. The first, which is most commonly used, is to schedule the appraisal of employees at a fixed interval, designating a time frame (one month, for example) within which all employees are evaluated. A second approach is to appraise each employee on every anniversary of the month the individual began work. Either of these two approaches is relatively simple administratively. A third approach is to schedule the evaluation at the completion of a task cycle. This may be the most feasible where tasks occur in clear, discrete sets (for example, the professor whose term begins in the fall and terminates in the spring). The final approach, which is least commonly used, is to vary the timing of the appraisal as a function of the task and the level of performance exhibited in the execution of duties. This latter approach suggests that evaluation should occur at the completion of the time period which is required for substandard performance to reveal itself. For example, if it takes six months to determine whether an individual is performing adequately, then evaluation should occur at the end of six months. Subsequently, evaluation should occur whenever it becomes evident that attempts to reconcile problems in performance are not forthcoming, or whenever new problems emerge. This is not to suggest that apparently acceptable performance need not be assessed; rather, it indicates that appraisal should occur more frequently among poor performers. While this approach makes intuitive sense, care must be taken not to apply unnecessary pressure on lower performers (which may decrease their motivation or be considered harassment). In addition, this type of approach will likely foster negative associations relating to performance appraisal, since evaluation is more likely to occur if "something has gone wrong".

---

[72] Steele, "In Praise of Appraisal", p.14.

Although we tend to think in terms of a time interval for appraisal, certain techniques, such as critical incidents, involve continuous monitoring of performance and recording of employee behaviour. This serves to remind us that, whatever interval is chosen, appraisal is, in effect, a continual process because evaluations are based on employee performance over time and feedback to the employee should occur frequently on a more informal basis. It is unrealistic to expect supervisors or subordinates to devote a great deal of time (and at frequent intervals) to the process of appraisal. Yet, informal monitoring and feedback can occur on a timely, day-to-day basis and then be formally summarized at evaluation time.

8.4 Appraisal Interviews

"No system yet devised deals with the core reasons for supervisor resistance to the performance appraisal process: we do not like to sit in judgement of our fellow man."[73] For this reason, the appraisal interview is often perceived as a sensitive issue which many supervisors consider the least palatable aspect of their job.

The keys to effective appraisal are careful preparation, open dialogue and two-way communication, and a problem-solving rather than fault-finding climate.[74] The literature is replete with advice and instruction on how to conduct an appraisal interview. Exhibit B-3 attempts to synthesize and distill the bulk of recent literature.

---

[73] Randall Brett and Alan J. Fredian, "Performance Appraisal: The System is not the Solution", Personnel Administrator, December 1981, p.61.

[74] Ibid., p. 62.

8.5 Linkages of Appraisal Systems to Other Personnel Systems

A system of performance appraisal does not exist in isolation from other human resource functions within an organization. In fact, it is appraisal's linkage with and contribution to other areas of the personnel system which makes it such an invaluable tool in human resource management. As figure B-1 illustrates, the information derived from an appraisal system is, or can be, linked to decisions for career planning and counselling; human resource planning and inventory systems; promotion, tenure and dehiring; compensation administration; and training and development.

8.5.1 Career planning and counselling

Social trends indicate that employees are increasingly desirous of taking a proactive stance in the management of their work and their careers. The developmental function of performance appraisal permits individuals to identify their own strengths and weaknesses and examine the fit of their needs and abilities with their specific function within an organization. Just as organizations utilize performance appraisal to maintain effectiveness and growth, so too does an individual look to his or her work to provide security, challenge, opportunities for self-development and a satisfying career. Data generated from performance appraisal aid in making career decisions by enabling employees to assess their suitability to and satisfaction with their job. In addition, the identification of problem areas in performance can lead to more effective career counselling and assistance, if the organization sees fit to make this service available to its employees. Career counselling and planning reduce the risk of personnel obsolescence, reduce turnover and personnel costs, and help both individuals and the organization to achieve their respective goals.

Exhibit B-3

Guidelines for Appraisal Interviews

Step 1: Prepare

- The employee should be contacted several days in advance to allow time to prepare.
- It should be a mutually convenient time.
- Assemble all data (documentation, appraisal instruments, etc.) and refamiliarize yourself with it.
- Ensure there will be privacy, lack of interruption, sufficient time.

Step 2: Outline the Purpose and Set the Tone

- Make subordinate comfortable.
- Get to the point of the interview (e.g., "As we discussed when we arranged this meeting, our purpose today is to discuss your performance since your last review and what we might do together to help you with any problems you may be having").
- Describe how you intend to conduct the meeting.
- Emphasize two-way communication.

Step 3: Invite Employee's Assessment of His/Her Performance

- Start the discussion by asking "How do you feel things are going on the job? What's going well and what problems are you having? How do you see your performance?" etc.
- If the employee had performance objectives, have your subordinate restate each objective and results obtained.
- Ask questions for purposes of clarification and elaboration.

Step 4: Listen and Learn

- Do not interrupt.
- Listen to what the employee is saying.
- Be prepared to alter your assessment in light of new data.

Step 5: Share Your Data

- Give specific examples of proficiencies and deficiencies.
- Emphasize employee strengths.
- Discuss performance, not personality.
- Tell employee what is wrong with their work, not themselves.
- Avoid comparing the employee to others.
- Allow the employee the opportunity to ask questions, clarify information.
- Explain why you assessed performance as you did.
- When discussing weak areas, focus on solving the problem (causes and possible remedies) rather than on the event.

Step 6: Summarize Your Joint Analysis

- Discuss various aspects of your data and the data of the employee; encourage mutual feedback.
- Praise the positive aspects of performance.
- Review what has been stated, and on what points you agree.
- Discuss where you differ, why you differ (remember complete agreement is not always mandatory) and how to resolve differences relating to future job performance.

Step 7: Develop a Plan of Action

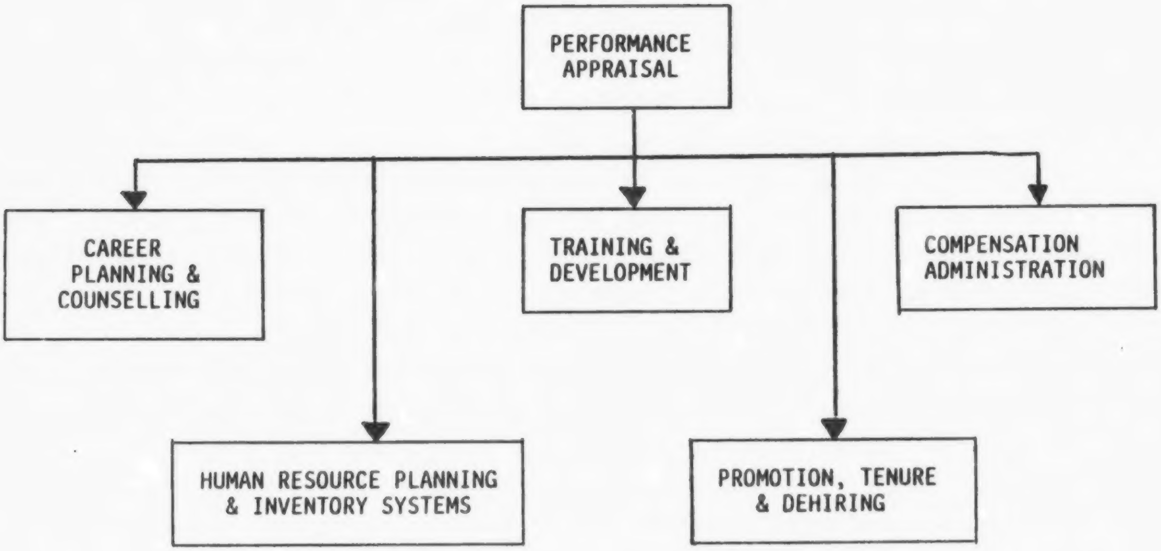
- Emphasize future improvement in performance.
- Encourage the employee to think out solutions to problems (e.g., What do you feel is the problem here? What do you think can be done? How? How can I help?).
- Suggest specific techniques to help improve performance.
- Set formal or informal performance targets and goals (process- or output-oriented).
- Outline jointly what changes are to be expected in the employee's performance in specific terms.
- Review action plans and set follow-up dates to review progress.
- End interview on a positive note of encouragement.

Step 8: Follow-Up

- Summarize the meeting in writing.
- Ensure employee has copy of appraisal.
- Meet with subordinate on agreed upon date to review progress.
- Make yourself available for coaching and counselling.
- Monitor employee for signs of improvement and given positive feedback.

Figure B-1

Linkages of Appraisal Systems to Other Personnel Systems


PERFORMANCE APPRAISAL
CAREER PLANNING & COUNSELLING
TRAINING & DEVELOPMENT
COMPENSATION ADMINISTRATION
HUMAN RESOURCE PLANNING & INVENTORY SYSTEMS
PROMOTION, TENURE & DEHIRING

8.5.2 Human resource planning and inventory systems

In the aggregate, performance appraisal can be used to assess the overall strengths and weaknesses of the organization's human resources and to identify the potential of employees in meeting future staffing requirements. In order to determine how many and what types of employees will be required in the future, it is necessary to compare forecasts of future needs with the present work force. Performance appraisal feeds data into the planning process about the capabilities and potential of the present work force for comparison purposes.

The basic tool for assessing the overall supply of people and talents available within the organization is the skills inventory.[75] Some of the information contained in these inventories may be obtained from the performance appraisal system. These inventories aid in a variety of organizational decisions: preparing long-range plans, forecasting future personnel needs, choosing individuals for special assignments, rotations or transfers, and assisting in the types of personnel decisions outlined below.

8.5.3 Promotion, tenure and dehiring

Performance evaluation systems are closely linked with promotion, tenure and dehiring decisions. Dehiring on the basis of incompetency or obsolescence of skills (as opposed to retirement or voluntary termination) can occur as the result of obtaining formal data from the performance appraisal system that an individual is not, nor is likely to be, capable of meeting the requirements of the position. In light of emerging legal and social trends regarding employee rights, it is particularly important that the appraisal process be valid and equitable in order to avoid lawsuits, grievances and charges of unfair or wrongful dismissal.

The implications of performance appraisal data for decisions on the granting of tenure are different from those concerning promotions. In the former case, the emphasis is on the perpetuation of good performance in the same position rather than on the prediction of good performance in a higher position. For promotion decisions, the issue is whether appraisal information should concentrate on success in the current job or on the potential and likelihood of success in the future job. A majority of appraisal systems only assess current job performance. If success on one job does not guarantee success on the next, then appraisal information may have to be augmented by mechanisms to evaluate employee potential.[76] Finally, the use of appraisal data for promotional decisions will be moderated by the extent to which merit and/or seniority are used as criteria for promotional purposes.[77]

---

[75] For more information, see Glueck, Personnel, p.116.

8.5.4 Compensation administration

Performance appraisal may or may not be closely linked to the compensation aspect of the personnel function. Although a BNA study in 1974 showed, for example, that 85 percent of American companies surveyed used performance appraisals for salary adjustments, Hildebrandt (1978) suggests that merit increases are often a more prominent notion in policy than in practice.[78] In point of fact, although existing theory justifiably advocates that employees can be differentiated on the basis of current performance, "many believe that seniority, age and therefore past performance and loyalty should have equal or greater weight in individual pay determination."[79]

In addition to this, the actual development of a fair and acceptable compensation system which succeeds in linking pay directly to performance is an awesome task for a number of reasons. First, individual differences in pay rates, it is argued, can be established on the basis of job performance, seniority, experience, professional qualifications or job category. Second, the individual may work in a job environment where work variances are beyond his or her control or where opportunities to effect improvement are limited. Third, individuals must be able to clearly perceive a direct link between their pay and performance, or the motivational aspects of higher pay will be severely reduced. Moreover, Marano (1979) proposes that

---

76. A relatively valid technique in the prediction of managerial potential is the "assessment centre". See section 7.4.3.

77. Deidre Gill, Performance Appraisal: Present Trends and the Next Decade (London, England: Institute of Personnel Management, 1977), p. 18.

78. F. Dean Hildebrandt Jr., "Individual Performance in Incentive Compensation", Compensation Review, 10 (1978), pp.28-33.

79. Glueck, Personnel, p.431.

organizations often resist using performance appraisals to allot wage adjustments due to their "legitimate insistence on controlling monies allocated to pay increases during a fiscal period, so as not to exceed the budget."[80] Despite these difficulties, compensation administration can still be effectively linked to performance appraisal.[81] Many organizations determine bonuses rather than wage increments, on data derived from performance appraisal. The challenge is to develop and implement a program sufficiently sophisticated in its design to be able to tie pay to performance in a way which overrides some of the aforementioned issues.

8.5.5 Training and development

Performance appraisal provides important data for input into the training and development function of the personnel system. The appraisal may highlight needs and opportunities for growth and development of the person, or the need to alter or update certain skills and knowledge; this helps the employee to function more effectively in his or her present job as well as to prepare the employee for possible future advancement. Growth and the acquisition of new skills and abilities may be accomplished by self-study, formal training courses, or job related activities, such as special assignments and job rotation. It is evident that training and development aid the individual as well as strengthening the organization by creating a better match between its human resources and the jobs necessary to meet organizational goals.

9.0 Results

9.1 Organizational Performance and Employee Satisfaction

The supervision of staff is oriented towards improving the performance level of subordinates or upgrading the quality of the employee's output. Alternatively, if the employee appears to be applying his or her full potential to a job for which he or she is appropriately suited, then supervision may involve no more than monitoring the maintenance of current levels of performance. Supervision, however, is usually more concerned with the former; staff development can be reasonably viewed as a continuing process in which the employee upgrades his or her skills and knowledge on an on-going basis as new information arises or as the demands of the job change. In this context, the intended result of a good performance appraisal program is improvement of the individual who is being evaluated. Collectively then, an overall program should enhance organizational performance.

---

[80] Richard A. Marano, "An Rx for Performance Appraisals", Personnel Journal, May 1979, p.306.

[81] The Hay System is an example of this approach.

This assumes that the performance expected of employees is congruent with the organization's goals. In the area of performance evaluation, a clear and acceptable articulation of the organization's goals and objectives is mandatory. Organizational goals determine the kinds of jobs an organization designs. These jobs and the specific tasks contained in them are determinants for the standards against which employee performance will be evaluated. If the overall goals and the process by which these are achieved are poorly defined, then the standards of excellence against which employees are appraised will also be poorly defined.

In the area of education, the results of performance appraisal should be to ensure that education personnel make the contributions to their students, their community and their society that are expected by these groups. In this way, performance evaluation is the essential building block of accountability. The final result of appraisal should be an improvement in the quality of the product; that is, the knowledge, competencies, attitudes, and potential of graduates. The intermediate results are those which ensue from the appraisal: improvement of employee performance and better educational decisions in areas such as placement, tenure, transfer, and promotion of educational personnel. In addition to these results, other outcomes may be expected to occur. Not only does performance appraisal serve to modify or change employee behaviour toward more effective working habits, it also provides a data base for research purposes or related personnel decisions. Furthermore, effective performance appraisal can foster supervisory understanding. Performance appraisal encourages supervisors to take an interest in their subordinates and offers the potential for greater understanding between the supervisor and his or her staff.

The end result of a good performance appraisal system should serve not only to improve employee performance but also to facilitate the attainment of the objectives of the appraisee, insofar as these do not conflict with the goals of the organization. The satisfaction of employees is not usually a criterion for evaluating the effectiveness of an appraisal program. There is in fact, a remarkably consistent body of literature that suggests that employee satisfaction is not a specific cause of high organization performance.[82] Nevertheless, job satisfaction is in no way antithetical to effective evaluation and can be considered a worthwhile purpose of performance appraisal. Indirect organizational outcomes may also evolve from a satisfied staff: more commitment among staff members, a better working climate, less absenteeism and turnover, and a better image in the community at large.

If an appraisal program is effective, regardless of whether employee satisfaction is an intent of the process, the result will be an increase in the level of employees' satisfaction with their jobs, provided that they are performing at an acceptable level. This type of reciprocity in employee satisfaction and positive organizational performance will be a natural outcome of a worthwhile appraisal system; that is, it will meet the needs not only of the organization as a whole, but also the needs of its individual members.

An effective appraisal system can also be seen as ensuring employee rights to fair treatment and due process. The open communication of performance expectations and the manner in which judgements about performance will be made should act to clarify for employees what it is they have to do in order to be successful in the organization. Secondly, employees receive a greater degree of protection from the effects of arbitrary decisions on performance by supervisors. Thirdly, the employee is protected by a system of due process if disputes arise about either performance criteria or performance judgements. In many systems, provision is made for employee retraining, employee assistance programs, or similar salvaging schemes so that problem employees are not summarily dismissed. Finally, an effective appraisal system ensures greater equity of compensation so that persons doing work of equal value in the organization have the right to equal compensation. None of these benefits are guaranteed by an appraisal system, but it should be obvious that a fair and valid system can only act to protect the rights of the capable performers in the organization and treat the less capable with all the due process to which they are entitled.

---

[82] See, for example, Edward E. Lawler, "Satisfaction and Behavior", in Richard M. Steers and Lyman W. Porter, eds., Motivation and Work Behavior (New York: McGraw-Hill, 1979), pp.287-301.

9.2 Maintenance of the System

A recent survey revealed that a distinct trend in appraisal practices in organizations is toward frequent change or revision of the appraisal system.[83] Maintenance of an appraisal system involves more than periodic assessments to determine that the appraisal process has remained unaltered from the time of its inception. New concepts are continually being developed in appraisal methods: people, organizations and the environment in which they operate are constantly changing. These changes place demands on program designers to be attentive to emergent problems or difficulties and to make modifications to the system as the need arises. Maintenance of the system necessitates a two-fold strategy: the first component of this strategy is a means to diagnose or evaluate the continuing effectiveness of the system; the second component is the establishment of a process for revision or renewal of the system in response to changing demands and the emergence of new problems.

The first step in diagnosis is to ensure that the objectives of the system are clear. In other words, yardsticks for measuring the success of the program depend on knowing what the specific purposes of appraisal are supposed to be and what results are expected. Once this has been established, it is necessary to determine where to look for evidence that these purposes are being served. Data may be obtained by examining the results of the appraisal process or examining the process of the appraisal itself.

An evaluation of the results of appraisal focuses on three areas: overall organizational effectiveness, individual performance levels, and satisfaction with the system. Depending on the type of organization, data on overall organization effectiveness can be elicited from company reports (profit, income, etc.), personnel records (absenteeism, turnover, grievances), from feedback from individuals external to the organization (clients, the community) and from monitoring changes in the organization's output (quantity, quality). Changes in individual performance levels will be forthcoming from the appraisal system, but can be supplemented by feedback from appraisers, appraisees, peers or outsiders (parents, clients, etc.) and, again, from personnel records. Satisfaction with the system is determined by verbal feedback from users of the system and by the extent to which the individuals involved are using or abusing the system in accordance with the procedures set down at the time of implementation.

---

[83] Teel, "Performance Appraisal", pp.296-316.

Diagnosing the process of appraisal itself will disclose shortcomings in the system and potential areas for improvement. Here, the focus is upon the degree of divergence from the design, techniques and implementation of the system established prior to diagnosis. It is likely that areas in need of improvement will be those in which the divergence has been most pronounced. In light of the voluminous studies which have concentrated on barriers to effective appraisal, it is useful at this point to outline some questions which illuminate the most frequently cited problem areas:

* Is the appraisal done at expected intervals on criteria previously established and fully understood by the employee?[84]
* Are policies and procedures complete, clear, consistent, and documented?
* Are accurate records being kept of appraisal judgements?
* Are appraisers given feedback on and held accountable for the quality of their appraisals?
* Is there sufficient opportunity for the appraisee to discuss the appraisal and introduce new comments or data into the record?
* Is there a channel of recourse for those who feel a particular appraisal is unfair?
* Does training of raters appear to be relevant and useful to the process of appraisal?
* Have there been complaints from appraisees of rater bias or lack of objectivity in the appraisal process?
* Is day-to-day supervisory behaviour encouraging more constructive and less threatening performance evaluations?[85]

---

[84] See Elaine R. Gruenfeld, Performance Appraisal: Promise and Peril (Ithaca, New York: New York State School of Industrial and Labour Relations, 1981), p.47.

[85] It has been found that supervisors who keep abreast of their subordinates' performance and provide fairly immediate feedback between appraisal interviews conduct better evaluations. (See, for example, Ronald J. Burke and Eugene Deszca, "Appraisal Interviews: Increasing their Effectiveness by Better Management of Day-to-Day Performance", Canadian Personnel and Industrial Relations Journal, May 1980, pp.18-31).

- * Are supervisors being expected to assess the performance of any tasks with which they are not personally familiar?
- * Are appraisal interviews being structured as a problem-solving or fault-finding exercise?
- * Are supervisors keeping a log on employee performance between appraisals or relying on memory alone -- or brief observation -- to conduct the appraisal?
- * Is the process of appraisal so time-consuming as to discourage thoroughness and effort among users?

Once the appraisal system has been diagnosed, the organization can revise weaker areas of the system, collect data on how well the changes work and then make further modifications. In effect the maintenance of the appraisal system is a reiterative process in which aspects of the system are modified or "fine-tuned" on an on-going basis. In order for renewal to occur, there must be a built-in channel for feedback from users of the system to those who are responsible for its design and modifications. These data must be coordinated and utilized to change the system as problems arise. Otherwise, the inflexibility of the system will create frustration among those for whom it has been designed and result in a loss of commitment to the system itself.

10.0 Conclusion

It is obvious from this review of the general literature on performance appraisal that it is a complex and controversial subject. A wide range of approaches to appraisal is possible. Very exacting expectations concerning what an appraisal system is supposed to accomplish can exist. Unfortunately, there is no one approach to appraisals nor methods of appraisals which can be said to satisfy all expectations or even to be superior to other approaches. Instead, the choice made from among the alternative approaches to appraisal seems to depend essentially upon the needs and objectives of the organization. For example, organizations historically opted for very simple approaches with expectations that might at best have been naive or the result of wishful thinking. However, such systems had the virtue of administrative simplicity and did not unduly occupy the time of supervisors or subordinates. Subordinates rated "satisfactory" or better would not feel inconvenienced by such systems. As long as the number of problem cases was small, such simple systems would be sufficient for organizations which placed a high priority on speed and simplicity of appraisal.

More recently, most organizations have had more complex priorities for appraisal systems and this has led directly to more complex appraisal systems. It seems to be inescapable that the more an organization wants from the appraisal system, the more complex the system has to be. To add to this dreary conclusion, is the effect of recent employee rights legislation and legal judgements which have placed a heavy legal emphasis on performance appraisal information to justify many personnel decisions. Thus most large organizations are required to develop more sophisticated appraisal systems whether or not they have modest ambitions for the uses of performance appraisal data.

Recent research and developments in performance appraisal have not been promising in terms of the practical needs of most organizations. Intensive research on behaviourally anchored rating scales (BARS) and behavioural observation rating scales (BORS) have produced rating methods which resemble psychological tests in their ability to satisfy psychometric criteria of reliability and validity. Unfortunately, these scales are highly job or situation specific and very cumbersome to develop. Their widespread use in an organization would require an extensive investment in preparation and continual monitoring to keep up to date with job changes. In addition, such techniques are essentially process-focused and do not look at results achieved from a person's work. In effect, the BARS and BORS approaches would be appropriate where there is a high degree of confidence that close conformity to established procedures produces the required results i.e., following standard operating procedure in a hospital ensures maintenance of hygiene standards. Whether anyone has sufficient confidence to rely solely on a standard set of teaching procedures for appraisal, BARS or BORS, to produce required results is very questionable.

Since different approaches to appraisal seem to have utility for specific situations or appraisal needs and no form of appraisal seems to be suitable for all purposes, an alternative perspective on appraisals might be considered. The alternative is to take a multifaceted approach to appraisals based on an identification of specific objectives for appraisals and different aspects of a job which might be suited to different forms of appraisal. Many organizations actually use multifaceted appraisals, but probably very few choose the different facets in a systematic manner to satisfy different requirements of the appraisal system. The concept of a multifaceted appraisal will be elaborated more fully in a later section of the report, but the following examples should illustrate the concept.

For entry level or probationary types of work, a candidate might need to possess an appropriate set of traits or abilities used in a particular work system and the demonstrated use of, or ability to learn the necessary skills. Suitable appraisal methods initially would use personality and skills criteria (assuming proper validation and reliability research), followed by behaviour criteria (assuming proper training of raters). As the person matures in their work and becomes proficient in their professional skills, a routine monitoring form of appraisal may be sufficient. For those who wish to or need to pursue professional development, a work planning and review (WPR)[86] or MBO approach would be appropriate. MBO and WPR approaches are also the most appropriate for those persons with administrative reponsibility or responsibility for the work of others. If the compensation system is essentially seniority-based and little or no provision exists for performance differentials in compensation, then an appraisal system is sufficient which distinguishes between "satisfactory or better" employees and "unsatisfactory" employees. If an organization plans to implement a major change in the style or methods of professional service, then an intensive, individualized WPR or MBO appraisal system would probably be required so that each individual would know in detail how they are supposed to work toward the change. These few examples should illustrate the concept of using an array of appraisal methods and criteria depending upon the situation of the employee and the needs of the organization. Fundamental to the multifaceted approach is the need for the organization to define its goals and priorities for the appraisal system and to recognize the realistic limits of some appraisal methods for some types of work. Underlying the use of any system of appraisal is the need to provide fair treatment and due process and the need to have validity, reliability, and acceptability for any method used.

---

[86] Work planning and review (WPR) is essentially the same as MBO, but with a more detailed specification of work methods to be used.

REFERENCES

Aahe, E. Lawrence Jr. "How do your Performance Appraisals Perform?". EEO Today, Autumn 1980, pp. 216-222.

Alewine, Thomas C. "Performance Appraisals and Performance Standards" Personnel Journal, March 1982, pp. 210-213.

Alpander, G. G. "Training First-Line Supervisors to Criticise Constructively". Personnel Journal, March 1980, pp. 216-21.

Baird, L. S.; Beatty, R. W.; and Schneier, C. E. The Performance Appraisal Sourcebook. Amherst, Mass.: Human Resource Development Press, 1982.

Beacham, S. T. "Managing Compensation and Performance Appraisal Under the Age Act". Management Review, January 1979, pp. 51-7.

Beaulieu, Rod. "An Easier Look at Performance Appraisal". Training and Development Journal, October 1980, pp., 56-58.

Beer, Michael and Ruh, Robert A. "Employee Growth Through Performance Management". Harvard Business Review, July-August 1976, pp. 59-66.

Bell, Robert R. "Evaluating Subordinates: How Subjective Are You?". S.A.M. Advanced Management Journal, Winter 1979, pp. 36-43.

Bernardin, H. J. "Effects of Rater Training on Leniency and Halo Errors in Student Ratings of Instructors". Journal of Applied Psychology, 63 (1978), pp. 301-308.

Bernardin, H. J. and Buckley, Ronald M. "Strategies in Rater Training". Academy of Management Review, April 1981, pp.205-211.

Bernardin, H. J. and Cain-Smith, Patricia. "A Clarification of Some Issues Regarding the Development and Use of Behaviorally Anchored Rating Scales (BARS)". Journal of Applied Psychology, August 1981, pp. 458-64.

Bernardin, H. J. and Pence, E. G. "The Effects of Rater Training: Creating New Response Sets and Decreasing Accuracy". Journal of Applied Psychology, 65 (1980), pp. 60-66.

Birch, William J. "Performance Appraisal: One Company's Experience". Personnel Journal, June 1981, pp. 456-460.

Blanz, R. and Ghiselli, E. E. "The Mixed Standard Scale: A New Rating System". Personnel Psychology, 25 (1972), pp. 185-199.

Borden, R. S. "Behaviorally Based Performance Appraisals". Internal Auditor, February 1980, pp. 36-43.

Bressler, R. B. and Lacy, A. W. "An Analysis of the Relative Job Progression of the Perceptibly Physically Handicapped". Academy of Management Journal, March 1980, pp. 132-43.

Brett, Randall and Fredian, Alan J. "Performance Appraisal: The System is not the Solution". Personnel Administration, December 1981, pp. 61-68.

Brinks, J. T. "Is there Merit in Merit Increases?". Personnel Administrator, May 1980, pp. 59-64.

Burke, Ronald J. and Deszca, Eugene. "Appraisal Interviews: Increasing Their Effectiveness by Better Management of Day-to-Day Job Performance". Canadian Personnel and Industrial Relations Journal, May 1980, pp. 18-31.

__________. "Managing Day-to-Day Job Performance: A Neglected Area of Training". Canadian Training Methods, April 1980, pp. 18-20.

Burke, Ronald J.; Deszca, Eugene; and MacDonald, Victor N. "Some Correlates of High Performance Indicators Among Managers in the Public Sector", in Proceedings of the Annual Conference of the Administrative Sciences Association of Canada, Organizational Behaviour Division, Ottawa, June 1982, pp. 59-68.

Burke, Ronald J.; Deszca, Eugene; and Weir, Tamara. "Using Feedback and Goal Setting to Improve P.A. Skills". Canadian Personnel and Industrial Relations Journal, October 1979, pp. 38-45.

Bush, Gerald W. and Stinson, John W. "Evaluating the Boss". Management Review, November 1980, pp. 14-17.

Buzzotta, V. R. and Lefton, Robert E. "How Healthy is Your Performance Appraisal System?". Personnel Administrator, August 1978, pp. 48-51.

Byham, William C. "Assessment Centres for Spotting Future Managers", Harvard Business Review July-August 1970, pp. 150-167.

Canada. Employment and Immigration Canada. Labour Market Development in the 1980's, Report of the Task Force on Labour Market Development. Ottawa: July 1981.

Cascio, Wayne F. and Bernardin, H. John. "Implications of Performance Appraisal Litigation for Personnel Decisions", Personnel Psychology, Summer 1981, pp. 211-227.

Cascio, Wayne F. and Phillips, N. F. "Performance Testing: A Rose Among Thorns". Personnel Psychology, Winter 1979, pp. 751-66.

Cheedle, M. Millar; Luthans, Fred and Otteman, Robert L. "A New Breakthrough for Performance Appraisal". Business Horizons, August 1976, pp. 66-73.

Cleveland, Jeanette N. and Landy, Frank J. "The Influence of Rater and Ratee Age on Two Performance Judgments". Personnel Psychology, Spring 1981, pp. 19-29.

Cummings, L. L. and Schwab, D. P. Performance in Organizations: Determinants and Appraisal. Glenview, Illinois: Scott, Foresman and Company, 1973.

Dartnell Corporation. How to Review and Evaluate Employee Performance. Chicago, Illinois: 1976.

Decotis, T. and Petit, A. "The Performance Appraisal Process: A Model and Some Testable Propositions". Academy of Management Review, July 1978, pp. 435-47.

De Nisi, Angelo S. and Stevens, George E. "Profiles of Performance, Performance Evaluations and Personnel Decisions". Academy of Management Journal, September 1981, pp. 592-602.

Dossett, Dennis L. and Greenberg, Carl I. Goal Setting and Performance Evaluation: An Attributional Analysis. Paper presented at the Annual Convention of the American Psychological Association, Montreal, Canada, September 1-5, 1980.

__________. "Goal Setting and Performance Evaluation: An Attributional Analysis". Academy of Management Journal, December 1981, pp. 767-779.

Dunnette, Marvin, ed. Handbook of Industrial and Organizational Psychology. Chicago, Illinois: Rand McNally & Co., 1976.

Etzioni, Amitai. Modern Organizations. Englewood Cliffs, N.J.: Prentice Hall, 1964.

Feeney, Edward J. "Developing the High Performance Edge". S.A.M. Advanced Management Journal, Autumn 1981, pp. 29-39.

Field, Lloyd M. "Performance Appraisal: Making It Meaningful". Canadian Personnel and Industrial Relations Journal, no. 3 (1979), pp. 15-20.

Forbes, R. J. and Anaya, E. M. "Appraiser and Appraised". Management Today, January 1980, pp. 33-34.

Fox, William M. "Consentient Merit Rating: A Critical Incident Approach". *Personnel*, July-August 1981, pp. 72-78.

Gabora, H. N. "An Effective Performance Appraisal". *Canadian Personnel and Industrial Relations Journal*, no. 4 (1977), pp. 25-32.

Gill, Deidre. *Performance Appraisal: Present Trends and the Next Decade*. London, England: Institute of Personnel Management, 1977.

Glueck, William F. *Personnel: A Diagnostic Approach*. Dallas, Texas: Business Publications Inc., 1978.

Gruenfeld, Elaine F. *Performance Appraisal: Promise and Peril*. Ithaca, New York: New York State School of Industrial and Labor Relations, 1981.

Halatin, T. J. "Evaluating the Superior Employee". *Supervisory Management*, December 1981, pp. 17-21.

Hale, J. "Communications Skills in Performance Appraisal". *Industrial Management*, March-April 1980, pp. 17-191.

Hamelink, J. and Hamelink, J. "A Numeric Plan for Performance Appraisal". *Training and Development Journal*, October 1980, pp. 80-90.

Harris, David. *Wrongful Dismissal*. Toronto, Ontario: Richard De Boo Limited, 1980.

Haynes, Marion. "Developing an Appraisal Program". *Personnel Journal*, January 1978, pp. 14-19.

__________. "Developing an Appraisal Program". *Personnel Journal*, February 1978, pp. 66-107.

Hays, S. W. "Human Resource Management: The Missing Link". *International Journal of Public Administration*, no. 3 (1980), pp. 297-330.

Heneman, Herbert G. "Comparisons of Self- and Superior Ratings of Managerial Performance". *Journal of Applied Psychology*, 59 (1974), pp. 638-642.

Heneman, Herbert G. and Schwab, Donald P. *Perspectives on Personnel/Human Resource Management*. Homewood, Illinois: Richard D. Irwin, 1978.

Henderson, Richard. *Performance Appraisal: Theory to Practice*. Reston, Virginia: Reston Publishing Company Inc., 1980.

Hildebrandt, F. Dean Jr. "Individual Performance In Incentive Compensation". *Compensation Review*, 10 (1978), pp. 28-33.

Hobson, Charles J; Mendel, Raymond M. and Gibson, Frederick W. "Clarifying Performance Appraisal Criteria". *Organizational Behavior and Human Performance*, October 1981, pp. 164-188.

Hyman, Earl M. and Elliott, Valerie. "Reactions to Performance Appraisal: The Stickor and Stuckee". *Canadian Training Methods*, February 1980, pp. 16-18.

Ilgen, Daniel R., et al. "Supervisor and Subordinate Reactions to Performance Appraisal Sessions". *Organizational Behavior and Human Performance*, December 1981, pp. 311-330.

Jaffee, C. L. and Cohen, S. L. "Improving Human Resource Effectiveness Through Assessment Centre Technology". In *Management of Human Resources*, edited by E. L. Miller, pp.350-379. Englewood Cliffs, New Jersey: Prentice-Hall, 1980.

Johnson, Robert G. *The Appraisal Interview Guide*. New York: AMACOM, 1979.

Kandel, W. L. and Langer, P. J. "Performance Evaluation and EEO". *Employee Relations Law Journal*, Autumn 1980, pp. 294-303.

Kerr, Steven, ed. *Organizational Behavior*. Columbus, Ohio: Grid Publishing Inc., 1979.

Kerzner, H. "Evaluation Techniques in Project Management". Journal of Systems Management, February 1980, pp. 10-19.

Kingstrom, Paul O. and Bass, Alan R. "A Critical Analysis of Studies Comparing Behaviorally Anchored Rating Scales (BARS) and other Rating Formats". Personnel Psychology, Summer 1981, pp. 263-289.

Kipnis, David; Schmidt, Stuart; Price, Karl; and Stitt, Christopher. "Why Do I Like Thee: Is It Your Performance or My Orders?". Journal of Applied Psychology, June 1981, pp. 324-38.

Kirby, Peter G. "Part 1: A Systematic Approach to Performance Appraisal". Management World, October 1981, pp. 16-28.

__________. "Part 2: A Systematic Approach to Performance Appraisal". Management World, November 1981, pp. 28-30.

__________. "Part 3: A Systematic Approach to Performance Appraisal". Management World, December 1981, pp. 28-44.

Kirkpatrick, Donald L. How to Improve Performance Through Appraisal and Coaching. New York: AMACOM, 1982.

Klimoski, R. J. and London, M. "The Role of the Rater in Performance Appraisal". Journal of Applied Psychology, 59 (1974), pp. 445-451.

Knowlton, W. A. Jr. and Ilgen, O.R. "Performance Attributional Effects on Feedback from Superiors". Organizational Behavior and Human Performance, June 1981, pp. 441-456.

Lacke, K. J., et al. "A Study of Employee Appraisal Systems of Major Cities in the United States". Public Personnel Management, March-April 1979, pp. 111-25.

Latham, G. P.; Wexley, K.N.; and Pursell, E. D. "Training Managers to Minimize Rating Errors in the Observation of Behavior". Journal of Applied Psychology, 60 (1975), pp. 550-555.

Lawler, Edward E. "Satisfaction and Behavior". In Motivation and Work Behavior, edited by Richard M. Steers and Lymon W. Porter, pp. 287-301. New York: McGraw Hill Book Company, 1979.

Lazer, Robert I. "Performance Appraisal: What Does the Future Hold?". Personnel Administrator, July 1980, pp. 69-73.

Lefton, R. E.; Buzzotta, V. R.; Sherbert, M.; and Karraker, D. L. Effective Motivation Through Performance Appraisal. Toronto, Ontario: John Wiley & Sons, 1977.

Levine, Edward L. "Let's Talk: Discussing Job Performance". Supervisory Management, October 1980, pp. 25-35.

Levinson, Harry. "Appraisal of What Performance?". Harvard Business Review, July-August 1976, pp. 30-36.

Letras, T. S. "A Practical Application of Peer Evaluation". Research Management, March 1980, pp. 13-19.

Love, Kevin G. "Comparison of Peer Assessment Methods: Reliability, Validity, Friendship Bias, and User Reaction". Journal of Applied Psychology, August 1981, pp. 451-457.

Lubben, G. L.; Thompson, D. E.; and Klasson, C. R. "Performance Appraisal: The Legal Implications of Title VII". Personnel, May-June 1980, pp. 11-21.

Mayer, Richard J. "Keys to Effective Appraisal". Management Review, June 1980, pp. 60-62.

McAfee, R. Bruce. "Performance Appraisal: Whose Function?. Personnel Journal, April 1981, pp. 298-299.

McConkie, D. L. "A Clarification of the Goal Setting and Appraisal Processes in M.B.O.". Academy of Management Review, January 1978, pp. 29-40.

McGuire, Peter J. "Why Performance Appraisals Fail". *Personnel Journal*, September 1980, pp. 744-746.

McMaster, John B. "Designing and Appraisal System that is Fair and Accurate". *Personnel Journal*, January 1979, pp. 38-40.

Miller, E. C. and Steinbrecher, D. O. "The Ten Commandments of 'Valid' Performance Appraisal". *Personnel*, September-October 1979, pp. 45-48.

Miner, John B. *The Management Process: Theory, Research and Practice*. New York: Macmillan Publishing Company Inc., 1978.

Miner, John B. and Miner, Mary Green. *Personnel and Industrial Relations*. New York: Macmillan Publishing Company Inc., 1977.

Morano, Richard A. "An Rx for Performance Appraisals". *Personnel Journal*, May 1979, pp. 306-307.

Moravec, Milan. "How Performance Appraisal Can Tie Communication to Productivity". *Personnel Administrator*, January 1981, pp. 51-54.

Murray, R. Stuart. "Managerial Perceptions of Two Appraisal Systems". *California Management Review*, Spring 1981, pp. 92-96.

Nix, Dan H. "Getting Ready for the Appraisal Interview". *Supervisory Management*, July 1980, pp. 2-9.

Patz, Alan L. "Performance Appraisal: Useful But Still Resisted". *Harvard Business Review*, May-June 1975, pp. 74-80.

Piamonte, John S. "An Employee Motivational System that Leads to Excellent Performance". *Personnel*, September-October 1980, pp. 55-66.

Quaglieri, Philip L. "Feedback of Feedback". Supervisory Management, January 1980, pp. 34-39.

Ragan, R. R. "Performance Appraisal: Setting the Pace for Your Employees". Management World, April 1980, pp. 8-10.

Saal, F. E. and Landy, F.L. "The Mixed Standard Rating Scale: An Evaluation". Organizational Behavior and Human Performance, 18 (1977), pp. 19-35.

Salloway, Roberta J. and Dayton, Allan S. "Personal Scorecards for the Overloaded Financial Officer". Management Focus, July-August 1981, pp. 28-32.

Sashkin, Marshall. "Appraising Appraisal: Ten Lessons from Research for Practice". Organizational Dynamics, Winter 1981, pp. 37-50.

Schick, Melvin E. "The 'Refined' Performance Evaluation Monitoring System: Best of Both Worlds". Personnel Journal, January 1980, pp. 47-50.

Schneier, Dena B. "The Impact of EEO Legislation on Performance Appraisals". Personnel, July-August 1978, pp. 24-34.

Schuler, Randall S. "Taking the Pain Out of Performance Appraisal Interviews". Supervisory Management, August 1981, pp. 8-13.

Sharzer, Lionel J. "Oh No, Not That Time Again!". Canadian Personnel and Industrial Relations Journal, no. 2 (1979), pp. 12-16.

Sikula, Andrew F. "Rethinking Present Appraisal Systems". Supervisory Management, March 1979, pp. 14-19.

Smircich, Linda and Chesser, R. J. "Superiors' and Subordinates' Perceptions of Performance: Beyond Disagreement". Academy of Management Journal, March 1981, pp. 198-205.

Smith, Howard P. and Brouwer, Paul J. Performance Appraisal and Human Development. Don Mills, Ontario: Addison-Wesley Publishing Company, 1977.

Smith, Michael. "Documenting Employee Performance". Supervisory Management, September 1979, pp. 31-37.

Smith, Patricia. "Behaviors, Results and Organizational Effectiveness: The Problem of Criteria". In Handbook of Industrial and Organizational Psychology, edited by Marvin Dunnette, pp. 745-775. Chicago, Illinois: Rand McNally and Co., 1976.

Spool, Mark D. "Training Programs for Observers of Behavior: A Review". Personnel Psychology, 31 (1978), pp. 853-888.

Steele, James W. "In Praise of Appraisal". S.A.M. Advanced Management Journal, Summer 1981, pp. 12-19.

Steers, Richard M. and Porter, Lyman W., eds. Motivation and Work Behavior. New York: McGraw-Hill Book Company, 1979.

Stewart, V. and Stewart, A. Practical Performance Appraisal. Farnborough, England: Gower Press, 1977.

Stinson, J. and Stokes, J. "Manpower: How to Multi-appraise". Management Today, June 1980, pp. 43-53.

Teel, Kenneth, S. "Performance Appraisal: Current Trends, Persistent Progress". Personnel Journal, April 1980, pp. 296-316.

Tosi, Henry L.; House, Robert J.; and Dunnette, Marvin D., eds. Managerial Motivation and Compensation. MSU Business Studies. East Lansing, Michigan: Michigan State University, 1972.

Truell, George F. Performance Appraisal: Current Issues and New Directions. Buffalo, New York: PAT Publications, 1980.

Villere, M. F.; Stearns, G. K.; and Lacko, K. J. "The Human Side of Performance Appraisal". Business, November-December 1979, pp. 46-48.

Wallance, Les. "Nonevaluative Approaches to Performance Appraisals". Supervisory Management, March 1978, pp. 2-9.

Walton, Richard E. "Work Innovations in the United States" Harvard Business Review, July-August 1979, pp. 88-98.

Wexley, Kenneth N. "Performance Appraisal and Feedback". In Organizational Behavior, edited by Steven Kerr, pp. 241-259. Columbus, Ohio: Grid Publishing Inc., 1979.

Winstanley, N. B. "How Accurate Are Performance Appraisals?". Personnel Administrator, August 1980, pp. 55-58.

__________. "Legal and Ethical Issues in Performance Appraisals", Harvard Business Review, November-December 1980, pp. 186-187.

Wooten, Bob. "Using Appraisals to Set Objectives". Supervisory Management, November 1981, pp. 30-34.

Yager, Ed. "A Critique of Performance Appraisal Systems". Personnel Journal, February 1981, pp. 129-133.

Zedeck, Sheldon. "Behaviorally Based Performance Appraisals". Aging and Work, Spring 1981, pp. 89-100.

C: TEACHER EVALUATION

The previous section of the report reviewed and assessed the state-of-the-art of performance appraisal in general. While most of the foregoing elements of our discussion have potential relevance to certificated educational staff, it is the purpose of this section to bring these elements into sharper focus. In particular, teacher appraisal will be examined: the need for an evaluation system, issues and components of evaluation, constraints on the process, and alternative approaches.

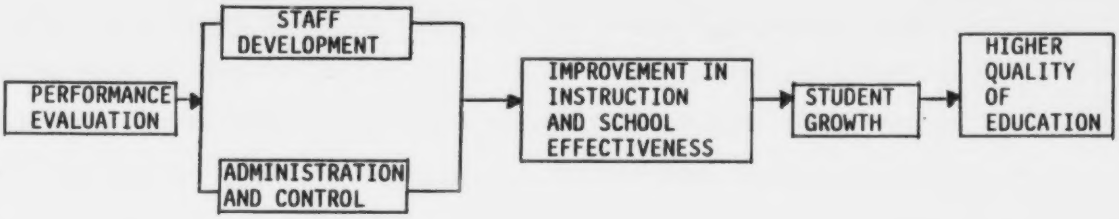
11.0 Need for an Evaluation System

In attempting to identify the reasons why any particular school requires an evaluation system, it is important to recognize at the outset that evaluation occurs on an on-going basis regardless of whether there is a system to formalize the process. Individuals form impressions and opinions of a teacher's behaviour and competencies whether or not a system exists to articulate or formalize these judgements. A teacher evaluation system, then, is needed in the first place to rationalize that process; in other words, to provide some order for the evaluation process and to introduce the components of responsibility, accuracy, and fairness.

At the same time that there appears to be a general consensus on the importance of teacher evaluation, responses to the question "Why do we need to evaluate?" are likely to vary considerably depending upon whom one asks. For example, a principal may see evaluation as a means to improve the overall effectiveness of the school, a teacher may expect greater opportunities for growth or freedom from arbitrary decision-making, and parents may wish for a better education for their child. A review of the current literature on teacher evaluation reveals a host of perspectives on the primary purpose(s) to be served by evaluation, but in reality, these reflect only variations in the emphasis given to the various results that a good evaluation system would be expected to achieve. That is, teacher performance evaluation may serve several needs simultaneously because the outcomes of the evaluation process are multifaceted. Figure C-1 illustrates this relationship.

Figure C-1

The Link Between Evaluation and the Quality of Education



PERFORMANCE EVALUATION
STAFF DEVELOPMENT
ADMINISTRATION AND CONTROL
IMPROVEMENT IN INSTRUCTION AND SCHOOL EFFECTIVENESS
STUDENT GROWTH
HIGHER QUALITY OF EDUCATION

Conceptually speaking, teacher performance evaluation describes a process which begins with a valid basis for initiating change in order to bring about behavioural improvement of education personnel. Staff development promotes the personal and professional growth of the teacher with the aim of improving instruction and providing the individual with feedback, direction, and motivation. Evaluation also serves as a basis for making administrative decisions relating to promotion, retention, transfers and/or compensation. It provides guidance to the teacher for the purposes of clarifying goals and objectives, ensuring acceptable performance, and optimizing the "fit" of the teacher with board policy and school procedures.

It is important to remember that although staff development, administration, and control are important outcomes of evaluation, they are not ends in themselves. As figure C-1 indicates, they represent intermediary steps in a process to bring about improvement in instruction and school effectiveness. These results in turn, serve to enhance student growth and achievement and hence, improve the quality of education in general. A good system of teacher evaluation serves a variety of needs within the educational hierarchy.

The atmosphere of accountability within which today's education personnel operate reflects increasing public demand for "quality education" at a "reasonable cost". These demands have contributed to the heightened interest in the evaluation of teachers as a means to control the educational process and to make improvements in the quality of instruction. In response to this growing accountability movement, the need for evaluation to serve an evidential role has become increasingly significant. There remains, however, the difficulty surrounding what, in fact, constitutes sufficient proof that teachers are performing at a level acceptable to concerned parents and society in general. As Holley and Hickman (1981) point out:

> ... the notion of "adequate evidence" or "adequate documentation" for, or of a given judgment concerning teacher competency may vary depending upon one's perspective or interest. The question "when has one adequately demonstrated the presence or absence of teacher competency?" may have more than one answer.[87]

Nor is the demand for evidence necessarily confined to those who are external to the field of education. Increasingly, evaluation is required to justify decisions to teacher unions or associations and various internal organizational members who seek clearly defined procedures for due process and the elimination of arbitrary administrative decisions.

While the concept of accountability is now in the forefront of educational thought, the notion loses meaning unless educators know how they are to be accountable and what evidence is necessary to satisfy those to whom they are in fact responsible. It is our contention that faith in the actual process of appraisal is critical (recall our analogy to the judicial system at the outset of this report) for two reasons. First, while the public may utilize elected representatives to protect their interests in the area of education, it is inconceivable that public access to data on individual teacher competence would be acceptable. Secondly, it is unlikely, at least in the short run, that a standard consensus of opinion will be reached on what constitutes explicit, acceptable proof of teacher competency and what the criteria of this evidence should be. It is therefore incumbent upon educators to develop a valid, fair, reliable and accurate process of appraisal which in and of itself satisfies the public expectation that an effective, ongoing effort is being made to improve the quality of instruction and enhance student growth.

In summary, teacher evaluation is needed for the following reasons:

- to rationalize the process of making evaluations;
- to promote personal and professional growth and, in so doing, to provide teachers with direction, motivation, feedback and opportunities to upgrade their knowledge and skills;
- to provide a valid basis for making administrative decisions relating to promotion, transfer, retention and/or compensation;
- to provide a method to monitor and control the quality of teaching and teacher-learner outcomes in accordance with school board expectations;
- to improve instruction and school effectiveness;

---

[87] Freda M. Holley and Randall C. Hickman. "Research on Teacher Evaluation: Needs and Realities". (Paper presented at the annual meeting of the American Educational Research Association, Los Angeles, April 1981), p. 6.

- to enhance student growth and achievement, both socially and academically;
- to improve the general quality of education in a way which is cost-effective;
- to provide evidence to educators, teacher associations and the general public that the goals and objectives of effective teaching are being achieved in a fair and systematic manner.

It should be noted that the first four points are concerned with the internal administration and supervision of the education process in a school. These purposes for teacher evaluation are very similar to the purposes for performance evaluation which might be found in any business or other professional service organization. The latter four points have more relevance to a public accountability requirement for teacher evaluation. These purposes do not have as many counterparts in other organizations. Business organizations are accountable in a much more general sense: Are they profitable? Do they run afoul of the legal system? The internal management processes of such organizations are rarely subject to much scrutiny as long as their respective publics are satisfied with the overall results of the organization. Public schools on the other hand are subject to considerable scrutiny of their internal processes. This is because the public is either concerned about overall results or because no satisfactory measures of overall results are available and internal processes must be scrutinized in lieu of something else to look at. In either case, there is an additional burden on teacher evaluation to satisfy a number of interested parties which is not found in conventional business or professional service organizations. The last four points then refer to the political aspects of teacher evaluation which are essentially unique to the public education system.

In light of present practices and the attendant controversies surrounding the means by which teacher evaluation should be implemented, it may seem idealistic at this point in time to expect an evaluation system to adequately serve such a variety of needs. In spite of an abundance of literature on teacher evaluation, there is a real dearth of empirical research on teacher evaluation methods. As Levin (1979) points out, "...for all the pages of writing there is not very much evidence about the reliability, validity or effects of most of the techniques used in teacher evaluation."[88] Nevertheless, there appears to be both a growing consensus about what kinds of fundamental components are necessary for effective evaluation. There are increasing numbers of examples of comprehensive evaluation methods that have been tested and applied in school settings.

---

[88] Benjy Levin. "Teacher Evaluation -- A Review of Research", Educational Leadership, December 1979, p. 240.

12.0 Critical Issues and Components

There are fundamental issues and concerns which must be addressed in order to develop a sound evaluation system. The subject of teacher evaluation has been historically fraught with conflicting opinions and judgements regarding its design, use and application. There are understandable reasons for the controversy. These will be elaborated in the subsequent sections. Our purpose now is to identify the salient issues of teacher evaluation and attempt to outline the characteristics or components of a good evaluation system.

12.1 The Criteria Question

Educators who are charged with developing a sound teacher evaluation system face their greatest challenge in trying to determine the appropriate standards or criteria against which to measure teacher effectiveness. Newton (1980) states:

> Education and our society need a system which defines with greater precision various levels of teaching performance in a way which not only allows accurate evaluation but also encourages improved performance.[89]

To accomplish this, the evaluation instrument must be able to identify the behaviours or characteristics that exemplify teacher competency.[90] There are basically three types of teacher criteria: presage criteria, output-oriented criteria, and process-oriented criteria.

12.1.1 Presage criteria

Presage criteria are the characteristics with which the teacher comes to the job, including personal and professional traits, training and intelligence. For example:

---

[89] Robert R. Newton. "Teacher Evaluation: Focus on Outcomes", Peabody Journal of Education, October 1980, p. 45.

[90] Section 6.1 of this report outlines the characteristics of effective criteria or performance standards, so they will not be repeated here.

Among those that have been found to be related to teaching effectiveness are scholarship in college, professional preparation, attitude toward teaching, and age and experience. Those found to have no significant relationship to teaching effectiveness include intelligence, knowledge of subject matter, cultural background, socio-economic status, sex and marital status.[91]

These criteria tend to be "removed in any demonstrable way from student achievement outcomes"[92] and often include teacher traits which usually make poor evaluation criteria. Moreover, presage criteria, with the exception of professional training, are generally more applicable for use in the recruitment and selection of teachers rather than as standards to be utilized on a continual basis over the teacher's career. For this reason, they do not ordinarily form part of the evaluation process.

12.1.2 Output-oriented criteria

Section 5.1 of this report summarized the major advantages and disadvantages associated with process-oriented and output-oriented systems of appraisal in general. In the area of teaching, process-oriented criteria are based on teacher behaviour and output-oriented criteria are based on student learnings.

Proponents of appraisal based on student learnings (output-based methods) advocate judging the relative effectiveness of the teacher by focusing primarily on pupil outputs and changes in pupil behaviour, rather than on teacher-learner processes. Advantages of this approach are "that it places the emphasis on results rather than intentions, forces more careful examination of pupil needs and related learnings, ensures a higher degree of pupil involvement and forces a closer review of teacher performance."[93] The major difficulty with these product-oriented criteria is demonstrating a direct causal relationship between pupil gain and the behaviour of a particular teacher. In addition: what the pupil brings to the classroom in terms of ability, motivation, previous knowledge, home and peer influence etc. is very powerful in determining academic standing at the end of the year; student achievement reflects only a portion of the total set of objectives for which a school may be held accountable; taking pupil standing at the year-end as an indicator of teacher effectiveness frequently does not recognize standing at the beginning of the year and the problem of adjusting for prior standing is extremely important and rarely recognized; this accountability system may reward the teacher who teaches to the text or who gives primary attention to those pupils who are below minimum standards but able to attain them; and such an evaluation system would probably reward teaching behaviour which promotes low cognitive level learning and penalize teaching which promotes complex learning.[94]

---

[91] Robin H. Farquhar, "How the Teaching Profession Measures Teaching Effectiveness". (Paper presented at the Conference on Teaching in Medical School, Saskatchewan, June 12, 1978), p. 5.

[92] Alexander Feldvebel, "Teacher Evaluation: Ingredients of a Credible Model", Clearing House, May 1980, p. 417.

[93] Robert V. Carlson and Richard Park, "Teacher Evaluation: Relevant Concepts and Related Procedures", June 1976, p. 6. (ERIC Document Reproduction Service No. ED 129 739).

12.1.3 Process-oriented criteria

Process-oriented criteria based on teacher behaviour emphasizes teaching methods and techniques, student-teacher relationships, and classroom strategies rather than pupil progress. This approach identifies broad areas such as planning, instruction, public relations, rapport with students, and communication. Within each of these areas, it delineates specifically expected behaviours or performance indicators. Strengths of this approach are that it forces a more clear definition of expected teacher performance, facilitates subsequent planning, and provides a framework for comparing performance among teachers in a system. Moreover, it is particularly conducive to the developmental aspect of evaluation. Carlson and Park (1976) cite the following as limitations of this approach: it may place an imbalanced emphasis on teacher behaviours rather than pupil outcomes; it may create high levels of conflict in attempting to resolve what teacher behaviours should be included; it may create inherent inequities when universally applied to all components of a school system; it may require a high level of staff time and involvement to fully define expected teacher behaviour; and generating a list of desired teacher behaviours may create unrealistic expectations.[95]

---

94. Robert S. Soar and Ruth M. Soar, Problems in Using Pupil Outcomes for Teacher Evaluation (Washington DC: National Education Association, 1975).

95. Carlson and Park, "Teacher Evaluation," p.10.

Appraisal against a prescribed or predetermined set of standards permits comparisons to be made more easily between teachers and between schools. These standards may take the form of an inventory of competency statements,[96] behavioural requirements, or performance indicators. One such list, developed by Richard P. Manatt of Iowa State University, is distilled from 360 items describing good teaching behaviours. It includes thirty items in five general categories to illustrate what Manatt calls the whole pool of valid, reliable and discriminating items. These are reproduced in exhibit C-1.[97]

If the major function of teaching is to enhance student learning then a whole range of criteria of teacher effectiveness may be considered. Despite considerable research into methods of breaking down teaching into a set of component competencies, no one has yet to identify a complete set of competencies that can be empirically demonstrated to be related to how much students learn. Yet the recent work of Manatt and others represents progress towards establishing some minimal set of valid criteria to be used as guidelines in the evaluation of teachers. In point of fact, chosen criteria may be partially a function of school policy and philosophy; that is, particular criteria are chosen or emphasized to reflect and embody the goals and mission of a particular school or district.

---

96. See, for example, Robert Knoop, An Inventory of Competency Statements for Teachers, Principals and Superintendents (Toronto: Ontario Ministry of Education, 1979).

97. Beatrice M. Gudridge, Teacher Competency: Problems and Solutions (Sacramento, California: American Association of School Administrators, 1980), p.32.

Exhibit C-1

Example of an Inventory of Competency Statements

Productive Teaching Techniques

1. The teacher uses probing questions for understanding of concepts and relationships and for feedback to the teacher.
2. The teacher uses student ideas in instruction.
3. The teacher uses structuring comments, such as examples to serve as "advance organizers".
4. The teacher uses varied teaching strategies and materials that stimulate student learning.
5. The teacher explains things well, puts ideas across logically and in an orderly way.
6. The teacher provides opportunities for pupils to learn material on which they will later be tested.

Positive Interpersonal Relations

1. The teacher shows respect for pupils.
2. The teacher is tolerant of students whose ideas differ from his/hers.
3. The teacher uses supportive criticism rather than blame, shame or sarcasm.
4. The teacher is readily available to students.
5. The teacher is fair, impartial and objective in treatment of pupils.
6. The teacher provides opportunities for all pupils to experience success.

Organized/Structured Class Management

1. The teacher constantly monitors pupil progress and adjusts the pace accordingly.
2. The teacher presents material in a well organized fashion in order to use class time efficiently.
3. The teacher has well defined objectives for pupils and works toward them.
4. The teacher uses pupil assignments that are relevant and sufficient for in-depth learning.
5. The teacher is businesslike and task oriented in behaviour.
6. The teacher keeps the "difficulty level of instruction" appropriate for each individual.

Intellectual Stimulation

1. The teacher inspires students to seek more knowledge.
2. The teacher is an exciting, vibrant person.
3. The teacher is enthusiastic.
4. The teacher sustains pupil attention and response with activities appropriate to the various pupil levels.
5. The teacher makes classwork interesting.
6. The teacher and pupils share in the enjoyment of humorous situations.

Desirable Out-of-Class Behaviour

1. The teacher is a good team worker.
2. The teacher strives for improvement through positive participation in professional growth activities.
3. The teacher assumes responsibilities outside the classroom as they relate to school.
4. The teacher is committed to the primary goal of assisting pupil growth.
5. The teacher utilizes community resources in instruction.
6. The teacher effectively reports pupil progress to parents.

Reproduced from Beatrice M. Gudridge, Teacher Competency: Problems and Solutions (Sacramento, California: American Association of School Administrators, 1980), p.37.

12.1.4 Discussion of criteria

To date there exists no unequivocal evidence to suggest that the "answer" to evaluation lies in utilizing only one of the previously discussed approaches. Neither can it be assumed that the optimum approach is a combination of the two types of criteria. Based upon the belief that evaluation tied to a single class of criteria has questionable validity, and in order to maximize the benefits of each set of criteria it is intuitively appealing to consider a hybrid approach to appraisal which combines several types of standards. However, comprehensive efforts in one school district to implement such an evaluation process (for which sufficient time has elapsed to judge the effectiveness of the program) resulted in failure. A multiple-criteria (process and product), multiple-appraiser approach was used. It is suggested that reasons for the failure of the system included hasty planning and implementation and an overwhelming amount of paperwork.[98] Therefore, although a combination of approaches may hold promise as a means of overcoming the limitations of a single approach, there is no empirical evidence to date to suggest that this is the optimum procedure. Nevertheless, as Feldvebel (1980) suggests:

> An evaluation system in which process and product criteria are utilized is probably a more valid predictor of teaching effectiveness than either of these factors alone. The evaluation of teacher effectiveness is a complex problem which should raise serious doubts about the validity of any narrowly classified judgment concerning the effectiveness of teachers.[99]

It is evident that criteria or standards of performance "should be explicit and as objective as possible and relate to those performance criteria that have a direct bearing upon job performance."[100]

It should also be noted that most sources concur that in defining what is expected of teachers, it is wise to involve teachers in the design, development or revision of criteria whenever possible; that is,

---

[98] See Richard R. Doremus, "What Ever Happened to Kalamazoo's Merit Pay Plan?", Phi Delta Kappan, February 1980, pp. 409-410; also William D. Coats Accountability in Education -- The Kalamazoo Plan presented at the Annual Meeting of the American Association of School Administrators, Dallas, Texas, February 1975.

[99] Feldvebel, "Teacher Evaluation", p. 418.

[100] Lawrence J. Tymko, "Performance Appraisal: At Issue in the 80's", Alberta School Trustee, November-December 1980, p. 10.

> ... administrators and staff members should be responsible for the selection of [teaching] criteria via a cooperative process. In fact, cooperative planning and joint effort hold the key to any successful evaluation program.[101]

Criteria determination involves more than just establishing validity and reliability; it depends upon the commitment and motivation of those upon whom these criteria will impact. Without the cooperation of teachers themselves, even ideally designed standards may prove inadequate.

12.2 Choice of Appraiser

In section 8.1 we considered five sources of appraisal: supervisory appraisal, peer appraisal, appraisal by subordinates, self-appraisal, and appraisal by outsiders. Reviewing the number of advantages and disadvantages identified with each source of appraisal, it is easy to understand the controversy surrounding who should be responsible for, or have input into, the process of teacher evaluation. While it has traditionally been the role of the principal to evaluate teachers based on classroom observation, the fact that this method is subject to bias and represents only a small portion of a teacher's total behavioural repertoire has led researchers and practitioners to investigate alternative sources of data.

12.2.1 Teacher self-evaluation

The difficulty with self-evaluation occurs in determining to what extent it should be part of the overall appraisal process. Allen (1970) states that "any change in teacher behaviour begins with a need to know yourself"[102] and, indeed, self-evaluation is perhaps the most valuable evaluation for the individual teacher, particularly because of the professional nature of the work and the difficulties of external ongoing observation of teaching activities. Those who claim, however, that self-evaluation alone is a sufficient mechanism for appraising performance fail to recognize that some form of supervision is necessary to ensure accountability to government, to district authorities and to society in general. And as Manatt (1980) points out, "[Self-evaluation] is not a very good predictor of teacher gain ...

---

[101] J. R. Marks, E. Stoops and J. King-Stoops, Handbook of Educational Supervision (Boston, Mass.: Allyn and Bacon Inc., 1978), p. 522.

[102] P. M. Allen, et al. Teacher Self-Appraisal: A Way of Looking Over Your Own Shoulder (Worthington, Ohio: Charles A. Jones, 1970), p. 22.

outstanding teachers tend to underrate themselves; the not-so-good, out of fear, overrate themselves."[103] Therefore, while self-evaluation is not sufficient by itself as a source of appraisal, it should be an integral part of the overall appraisal program.

12.2.2 Multiple appraisers

A recurring suggestion has been to use a team approach to improve teacher evaluation. The team might include the principal, the department head, the teacher and, at times, the area superintendent. On the one hand, a team approach can be more intimidating to the teacher, and more costly and time-consuming in general. On the other hand, there is evidence that the use of multiple appraisers in any context tends to increase the validity and reliability of the appraisal process and reduce errors of bias and subjectivity. The validity and reliability of this approach then must be weighed against its acceptability to teachers, its cost, and the added time and effort required to conduct the appraisal.

12.2.3 Peer appraisal

There are a number of suggested approaches for incorporating peer appraisal into the overall teacher evaluation system. Teachers within the same department could visit each other through informal invitations or a planned approach. Mallery (1975) describes a program in one Massachussetts school in which "an evaluation team consisting of the head of the department, a colleague in the same department and a colleague from a different department ... visit a class for three consecutive days, write a report and then share the information with the evaluatee."[104] Alternatively Carlson and Park (1976) suggest that "one plausible approach might be for a teacher to nominate potential persons best suited to judge a teacher's performance, and the supervisor pick a predetermined number of persons at random from the proposed list of names."[105] There is little doubt that fellow professionals may provide valuable input for the assessment of teacher performance. Moreover, peer teachers are "more apt to take into account day-to-day classroom demands and circumstances than are principals."[106] Yet there are distinct limitations to the usefulness of peer appraisal. First, as discussed earlier, peer appraisal is not highly recommended in organizations where peers compete for similar organizational rewards or use of resources. Second, timetable restrictions may impede the practicability of this approach. Third, subjectivity and the propensity for bias remain potential problems. Finally, many teachers strongly oppose peer evaluation. As Haefele (1980) states:

---

[103] Quoted in Beatrice M. Gudridge, Teacher Competency: Problems & Solutions (Sacramento, California: American Association of School Administrators, 1980), p. 38.

[104] D. Mallery, The Strengths of a Good Faculty: Notes on Evaluation, Growth and Professional Partnership of Teachers (Boston, Mass.: National Association of Independent Schools, 1975), p.35.

[105] Carlson and Park, "Teacher Evaluation", p.15.

> Negative peer reports ... could create disharmony and alienation among faculty. Their [the teacher's] perceptions of teachers and administrators as distinct groups with different functions and roles prompt them to assign teacher evaluation responsibility to administrators.[107]

12.2.4 Appraisal by outside experts

Many teachers speak of the possibility of appraisal by expert individuals, such as resource teachers, as a means to reduce subjectivity and obtain valuable feedback from informed individuals with considerable teaching expertise. In reviewing the pros and cons of this approach, it would appear that trained resource people may best be utilized on an intermittent or supplemental basis, in situations where teachers are experiencing special problems or difficulties which cannot be readily resolved by the teacher alone or with assistance from his or her supervisor.[108]

12.2.5 Appraisal by students

Research in this area has focused primarily on postsecondary institutions and findings have been inconclusive. For example, Levin (1979) states that "the most difficult problem with students' evaluations of teachers has to do with their validity;"[109] on the other hand,

---

106. Donald L. Haefele, "How to Evaluate Thee, Teacher -- Let Me Count the Ways", Phi Delta Kappan, January 1980, p. 350.

107. Ibid.

108. See section 8.1.

109. Levin, "Teacher Evaluation", p.241.

Sandefur's findings conclude that: "pupil evaluation of teachers is a reasonably valid source of information."[110] Hence, a great deal of controversy surrounds the use of students as evaluators. In brief, the following evidence seems to be indicated:

* The general consensus[111] is that several student rating tools are reliable.[112]

* Teachers who have a high degree of interaction with students are rated high by students, but they also receive ratings (from principals and others) that do not correlate well with the student ratings.[113]

* Student characteristics such as age and sex are not related to student evaluation results.[114]

* Class size, teacher reputation, student interest, and grades do tend to influence student ratings.[115]

* Clarity of presentation, enthusiasm and empathy with students appear repeatedly as desirable characteristics of good teachers[116] in student evaluations.

* Student evaluations are discriminating to the extent that they form a bigger sample and offer many more observations from which to make an overall evaluation.

* In general research indicates a significant, albeit modest, correlation between student achievement and student ratings of teachers.[117]

Students' ratings, like peer appraisal or appraisal by outside experts, cannot be utilized as a sole or major component in the evaluation of teachers. Student appraisal can make a significant contribution to a teacher's development because of the valuable feedback

---

110. J. T. Sandefur, An Illustrated Mode for the Evaluation of Teacher Education (Washington, D.C.: American Association of Colleges for Teacher Education, 1978), p.20.

111. See Irvin Lehman, "Evaluating Teaching" in William Gephart [ed.], The Evaluation of Teaching, (Bloomington, Indiana: Phi Delta Kappa Centre on Evaluation and Research, 1974).

112. Recall that reliable means different groups of students tend to give similar ratings to the same teachers, and the same students will rate teachers similarly at two different times.

113. David Rosevear "Teacher Evaluation: Teachniques and Tools", Comment on Education, February 1978, p. 16.

114. Janice Smith and T. J. Brown, "Relationship Among Secondary Students' Evaluations of Their Courses and Teachers and Their General Attitude Towards School", 1976 (ERIC Document Reproduction Service No. ED 129 867).

115. Levin "Teacher Evaluation", p.241.

116. Ibid.

117. David Potter; Paul Nalin; and Ann Lewardeski, The Relation of Student Achievement and Student Ratings of Teachers (Princeton, New Jersey: Educational Testing Service, 1973).

which it provides. However, although this source of appraisal makes a worthy contribution to the assessment of teaching in many university and community college settings, the age and variations in development stages among public school students precludes the standardized use of student appraisal for evaluation purposes in the public school system.

12.2.6 Discussion of appraisers

In determining who should conduct the appraisal, it is best to consider the alternatives as potential sources of information or data on teacher performance. As such, a good evaluation is one which is a balanced assessment that draws upon numerous samples of behaviour and numerous sources of information. As Carlson and Park (1976) point out:

> To rely on one person making only a few observations over a long period of time can only provide a very limited view of a person's performance, regardless of how diligent or sincere the motives of such a single observer.[118]

Pragmatically, it may be implausible or cumbersome to consider all sources, but it should be remembered that the more numerous and varied the sources of information, the greater the potential for increasing the validity, reliability, and objectivity of the assessment.

12.3 The Role of Goal Setting

The concept of performance objectives in education has evolved over the last two decades. This approach involves a dialogue between a supervisor and supervisee who mutually develop goals and objectives for an upcoming period of time. Unlike measurement against a predetermined set of criteria, performance is assessed against mutually-derived goals and objectives. A strong proponent of this approach, George Redfern (1980), claims that "evaluation actually should be regarded as a diagnostic process, enabling individuals and their evaluators to focus on appropriate objectives -- objectives that, if accomplished, will produce better and more effective services."[119]

---

[118] Carlson and Park, "Teacher Evaluation", p.15.

[119] George B. Redfern, Evaluating Teachers and Administrators: Performance Objectives Approach (Boulder, Colorado: Westview Press, 1980), p.8.

The process of setting goals and objectives to improve individual performance has been the subject of considerable study in management and behavioural science fields, particularly the effects of goals and feedback on learning as well as motivation. In an extensive review of research on the application of goal setting, Latham and Yukl (1975) concluded:

> As for the practical feasibility of goal setting as a means of improving employee performance, the research shows goal setting programs to be effective over an extended time period in a variety of organizations, at both managerial and non managerial levels.[120]

Furthermore, there is strong support for propositions that "specific goals increase performance and that difficult goals, if accepted, result in better performance than do easy goals."[121]

Unlike traditional classroom observation techniques where the teacher plays a relatively passive role in the evaluation process, goal setting provides the teacher with the opportunity to participate actively in establishing the expectations against which he or she will be assessed and in defining skills or activities necessary for future behavioural improvement.

Many of the advantages and disadvantages of management by objectives (MBO) outlined earlier[122] apply to the concept of goal setting and performance objectives in an educational environment. In addition, among its strengths for teacher evaluation, are also the following:

* The method clarifies teacher and supervisor roles and removes ambiguity concerning respective areas of responsibility;

* The method encourages cooperation and teamwork between the teacher and the supervisor;

* An opportunity is provided for meaningful teacher input and the identification of special or problematic circumstances in the teaching environment;

---

[120] Gary P. Latham and Gary A. Yukl, "A Review of Research in the Application of Goal Setting in Organizations" in Richard M. Steers and Lyman W. Porter, eds. Motivation and Work Behavior (New York: McGraw Hill Book Company), p.505.

[121] Ibid., p. 504.

[122] See section 7.4.2.

- Goal setting tends to be a motivational tool in improving performance;
- There is less tendency for the personality of the teacher to become an issue; and
- The process encourages and supports self-evaluation.

On the other hand, some of the major disadvantages are:

- The method is time-consuming and requires considerable commitment from all participants to the process;
- It is often difficult for teachers to identify and formulate realistic job targets;
- Special skills and training are required by both the supervisor and the teacher;
- The method requires a relatively long time frame to ascertain student gains, yet long time spans are not suitable for effective feedback;[123]
- While this approach is a useful developmental tool, it may be difficult to translate information derived from the process into data for administrative and control purposes.

Because objectives are derived individually by supervisor/subordinate dyads, goal setting introduces a strong element of flexibility into an appraisal program in order to accommodate individual differences and differences in terms of teaching style, environmental demands, curriculum guidelines, and availability of physical facilities. It also avoids the complex problem of trying to generate a complete set of prescribed standards which are at the same time both specific and universally applicable. Moreover, teacher resistance is often reduced by the active participation of teachers in the process itself. For goal setting to be effective, however, participants must be committed to the process and willing to expend the time and effort required to make it work.

It should also be noted that although MBO and goal-setting approaches are typically used in output-oriented appraisal systems, this does not have to be the case. Many of the goal-setting systems in teacher evaluation actually use process goals more than output goals. A process goal is concerned with planning and goal setting in order to use a particular teaching method, or to implement changes in the teaching process. For example, a process goal might be a statement of intent by a teacher to elicit greater participation and discussion in the classroom, or the use of more student research projects to learn certain concepts.

---

[123] Education Research Services, Evaluating Teacher Performance (Arlington, Virginia: Educational Research Services Inc., 1978), p.7.

Evaluations of the achievement of these objectives might be through classroom observations by a principal focusing on the comparative frequency of one-way vs. two-way interaction in a class period, or a count of the number of student projects in one term compared to another term. If the principal and the teacher both agree on the merits or value of using certain teaching methods, then it is quite valid to use goal setting for the adoption of these methods in operational terms and to be assessed on these criteria.

12.4 Formative and Summative Evaluation

Educators frequently make the distinction in appraisal between "formative" and "summative" evaluation:

> Formative teacher evaluation helps teachers improve their performance by providing data, judgments and suggestions that have implications for what to teach and how. On the other hand, summative teacher evaluation serves administrative decision-making with respect to hiring and firing, promotion and tenure, assignments, and salary.[124]

In accordance with this distinction, there is usually the attendant assumption that the two categories of evaluation are in direct conflict with one another, that is, teacher evaluation for the purpose of promoting growth and development is at odds with evaluation for the purpose of making personnel and administrative decisions.

From the supervisor's point of view, this twofold purpose of appraisal often gives rise to what supervisors perceive as a "judge-helper paradox" in their role.[125] Similarly, teachers' level of resistance to evaluation is dependent upon which purpose evaluation serves. Research studies[126] suggest that "teachers have negative feelings when they see the evaluation as serving administrative purposes. Conversely, teachers who believe the evaluation is for instructional purposes exhibit more positive attitudes."[127]

---

[124] Jason Millman, ed. Handbook of Teacher Evaluation (Beverly Hills, California: Sage Publications, 1981) p. 13.

[125] See section 4.0.

[126] See, for example, M. J. Zelenak and B. C. Snyder, "Teacher Perceptions of the Teacher Evaluation Process", California Journal of Education Research, 25 (1974), pp. 205-212 or B. King and W. Jordan, "Evaluation and Assessment of the Performance of Certificated Personnel" (Fresno, California: Fresno County Department of Education, July 1972). (ERIC Document Reproduction Service No. ED 104 907).

Perhaps some degree of conflict between formative and summative evaluation is inevitable, but it is important to note that the level of conflict is directly proportional to the quality and effectiveness of the evaluation system itself. It is the intention of a good evaluation system to provide methods of appraisal and standards of performance which are valid and equitable. This not only helps teachers identify ways to improve their performance, but also reduces the likelihood that the administrative decisions which affect them will be arbitrary or unfair. Furthermore, "administrative judgments are often instrumental in helping teachers develop their skills and conversely, the developing of individual skill serves the purpose for which institutional judgments are intended."[128] A good evaluation system, one which is valid, reliable and fair, is in the best interests of good teachers. Although it may not be appreciated by poor teachers who, with repeated assistance and developmental effort, are unable or unwilling to improve.

A currently unresolved dilemma related to summative evaluation is the issue of control. At the classroom level, the determination of the purpose(s) of appraisal comes down to creating a balance between the need for control to ensure that the school board program and policy objectives are being met and the need for teacher autonomy. On the one hand, the teaching profession requires a monitoring and control process to ensure that the goals of education are being adequately achieved, and to correct any serious deviations from expected standards that may be revealed by the performance appraisal instrument. On the other hand, the professional nature of teaching necessitates a high degree of autonomy and flexibility on the part of teachers. There exists, therefore, the complicated problem of distinguishing flexible appraisal practices from laissez-faire neglect and appropriately controlled performance from demotivating or counter-productive constraint. In the absence of any evidence to suggest that self-evaluation by itself is a sufficient mechanism to control the quality of teaching personnel, the control aspect of performance evaluation must remain. The question of where to draw the line on the continuum between imposed and self-control is clearly a matter of opinion. The emphasis, therefore, should be on minimizing the adverse or dysfunctional effects of control. A number of factors can contribute to a reduction of adverse effects:

---

[127] Jean Goodwin, "Constraints on Teacher Education" (Dawson, Georgia: Terrell County Board of Education, 1977), p.24. (ERIC Document Reproduction Service No. ED 147 299).

[128] James O'Hanlon and Lynn Mortensen, "Making Teacher Evaluation Work", Journal of Higher Education, November-December 1980, p. 664.

- The information which an appraisal system generates should be accurate and timely;
- The standards set by the appraisal instrument should be viewed as reasonably attainable;
- The purpose(s) of the evaluation should be made explicit to those involved in the process;
- There should be some system of recourse or appeal against controls which are perceived to be inequitable;
- Job responsibilities should be fixed to encourage cooperation rather than competition;
- The evaluation system should be understandable to its users;
- There should be participation in developing and maintaining controls of the system (e.g., mutually-derived goals); and
- The evaluatee should be monitored on those items which fully reflect behaviours that are under the direct control of the person.

It may not be possible to extinguish the potential conflict between summative and formative evaluation. From both the supervisor's and the teacher's perspective, a certain amount of discomfort may, in some instances, be associated with evaluation for purposes of administration and control. An effective evaluation system, however, comes closest to creating a balance which recognizes the need for both, by attempting to base decisions and control processes on systematic, objective and justifiable assessments of individual performance.

Trade-offs in emphasis between summative and formative evaluation can occur depending upon the needs or situation of a particular education system. If one school district has a goal of implementing a new philosophy of education with attendant changes in teaching methods, then close attention to formative evaluation would be required. If another district has great confidence in the professional abilities of the teaching staff, then formative evaluation is not as necessary and summative evaluation may be sufficient. If some problems of teacher competence are determined to exist in some location, then both formative and summative evaluation will be used to try to salvage or correct the performance problem. Thus there are no easy rules for either justifying or ruling out either formative or summative approaches to teacher evaluation. The most important principle is to determine what needs or purposes are to be served by the evaluation system and then choose whichever combinations of evaluation approaches best satisfy those needs.

12.5 Components of a Good Evaluation System

Current literature reveals a variety of approaches to teacher evaluation, and these may differ significantly from one school to another. In spite of these pronounced variations, there are certain components or characteristics which are a part of any good evaluation system or which reflect, in the absence of perfected application, the intentions of the system. These are summarized in the twenty-five points that follow.

1. Every effective system of personnel evaluation begins with the development of a philosophy of education and the purposes which will be served by evaluation of teacher performance. The philosophy and purposes are the rationale upon which the entire evaluation program is based.[129]
2. The processes and bases (i.e., the variables being observed or rated) must be made explicit and must be understood by all the teachers and administrators in the system.[130]
3. Criteria or standards should be based on the employee's actual job responsibilities, jointly developed goals (if applicable), program objectives and instructional priorities as they relate to student learnings.
4. Evaluation must be a cooperative endeavour, with the full support, involvement and commitment of teachers, and designed to establish rapport and communication between evaluators and evaluatees.
5. Evaluation should be systematic, verifiable, specific, objective, and cost effective.[131]
6. Teachers should be held accountable only for those variables which are under their control.[132]
7. Formal evaluation must be augmented by day-to-day feedback, monitoring and support by supervisory personnel.
8. A system must be available for appealing those evaluations which an evaluatee perceives as unjust.

---

[129] Harold E. Landrith, "Systems Approach to Evaluation". (Paper presented at the American Association of Colleges for Teacher Education Conference, Chicago Illinois, March 1977), p.2.

[130] Donovan Peterson and Annie Ward, eds., Due Process in Teacher Evaluation (Washingston, D.C: University Press of America Inc., 1980), p.157.

[131] William R. Tracey, "Teacher Evaluation: Another Perspective", Clearing House, January 1978, p. 241.

[132] Peterson and Ward, Due Process, p.158.

9. Unless otherwise agreed upon by all staff members of a particular school, the data generated from an evaluation should be regarded as confidential.

10. The prerequisite of any good evaluation program is a clear and comprehensive definition of the duties and responsibilities of each position.[133]

11. Evaluation should be designed with sufficient flexibility to be successfully adapted to a broad range of educational circumstances and to provide for differences in the context of staff members' assignments.[134]

12. Evaluation must be planned. The institution or revision of an evaluation system requires a long time frame for completion and the extended efforts of staff at several different levels of the educational hierarchy.

13. The impetus for designing or revising a program should come from the highest level and have the full support of the school board, the province, and teacher associations.

14. Systematic evaluation should be viewed as only one means of stimulating improvement. Other means are helpful supervision, a supportive climate, and optimum quantities of learning materials and supplies.[135]

15. Those responsible for evaluation should be trained in the use of the system and provided with periodic opportunities for additional evaluation training.

16. Many non-classroom activities are important to the total development of the students in the school. Any activity which contributes positively to the development of the total school program should be valued and become part of the evaluation system.[136]

17. The purposes (formative, summative etc.) to be served by the evaluation program must be clearly articulated and understandable to all staff.

18. Evaluation should be tailored to fit the needs of the local school division,[137] in order to be congruent with the division's instructional goals and curriculum philosophy.

19. Documentation is a fundamental product of an evaluation system. These documents should be meaningful, clear and relevant; they should not be so complex or time-consuming to prepare and read as to discourage their use. All evaluations should be communicated in writing to the persons being evaluated.

20. Follow-up support should be provided for the improvement of performance noted in evaluation reports. It should include a variety of sources such as in-service education, observation opportunities, peer assistance, and aid from external resource people, additional or differing materials.[138]

---

[133] Redfern, Evaluating Teachers, p.2.

[134] Craig J. Melvin, "Performance Expectations: Teacher Supervision and Evaluation", Saskatchewan Education Administrator, December 1980, p. 11.

[135] Redfern, Evaluating Teachers, p.4.

[136] Peterson and Ward, Due Process, p.158.

[137] Robert J. Berardi, "Evaluation of Educational Personnel in the Commonwealth of Massachusetts", in Robert J. Munnelly ed., "The Improvement of Local Staff Evaluation and Supervision (Massachusetts Association For Supervision and Curriculum Development, 1981), p.2. (ERIC Document Reproduction Service No. ED 208 525).

[138] Ibid., p. 3.

21. Evaluation should not obstruct freedom for teacher initiative in classroom experimentation of methods, nor interfere with a classroom climate of satisfaction, creativity and accomplishment.[139]

22. The evaluation system should be logically related to some method (rewards, compensation, recognition, support) which acknowledges and encourages growth or improvement and stimulates even greater accomplishments.[140]

23. The evaluation system should be realistic in terms of time and funds for implementation.

24. The evaluation process should be a problem-solving rather than fault-finding exercise. It should emphasize development and generate specific suggestions on how to improve performance.

25. The evaluation system should have a built-in strategy for monitoring the effectiveness of the overall system and revising problem areas.

The preceding characteristics of a good teacher evaluation system can serve as guidelines in assessing the approaches to be discussed. Before examining these approaches, however, it is useful to look at the major factors which constrain or complicate the process of teacher evaluation. Current literature on performance appraisal in general is replete with testimonials on the problems ensuant to system design and application. The following section focuses on factors specific to the area of teaching, factors which in many cases serve to demonstrate why some areas of teacher evaluation remain controversial.

13.0 Constraints on the Evaluation Process

There are a number of general factors particular to the area of teaching which tend to constrain the development and implementation of an effective teacher evaluation process. These factors and their implications for evaluation are outlined below.

13.1 Lack of Validated Research

Although the concept of teacher evaluation has received a great deal of attention, particularly in recent years, little of this attention has been in a form which generates valid empirical data applicable to teaching in general. As Levin (1979) states, "the entire field of teacher evaluation has suffered from a surplus of opinion and a shortage of evidence".[141]

---

[139] Marks, Stoops, and King-Stoops, Handbook of Educational Supervision, p.518.

[140] See, for example, Gudridge, Teacher Competency, pp.21-23.

[141] Levin, "Teacher Evaluation", p.244.

This means that designers of evaluation systems lack a solid substratum of systematic and well-researched information upon which to build newer or better evaluation systems.

Implications

In spite of this constraint, designers can expect a proliferation of new research in the future. Holley and Hickman (1981) claim that certain developments (such as the publicity and public concern about teacher competence and the rising interest of teachers associations in evaluation to ensure fairness are creating a more favourable climate for research on teacher evaluation.[142] Additional impetus is also coming from the fact that current practices in teacher evaluation are, by and large, in need of major improvement. In the meantime, designers have the opportunity to draw valuable insights from actual examples where teacher evaluation systems have been successfully applied. Particularly in the United States, where many states have legislated comprehensive teacher evaluation programs, we see applications which are on the cutting edge of new and innovative appraisal practices.[143]

13.2 Teacher Resistance

Hammons (1975) attributes negative teacher attitudes to a general disillusionment with traditional evaluation systems.[144] In addition, many teachers find evaluation threatening or may be distrustful of the process. Anxiety occurs because evaluation is used to make administrative decisions or to control the teaching process in a way antithetical to the teachers' styles, methods and techniques. Distrust arises when teachers question the use, validity, and objectivity of the appraisal process. Extreme anxiety may cause performance to be unrepresentative of the teacher's best efforts. Furthermore, if teachers are anxious or distrustful, resistance or demotivation may cause a deterioration in performance and quality of instruction.

---

[142] Holley, and Hickman, "Research on Teacher Evaluation", p.3.

[143] See the field study report (volume 2) of this project.

[144] A. L. Hammons, "How to Evaluate Teachers on Performance". (Paper presented at the annual meeting of the National School Boards Association, April 1975), p. 2.

Implications

This would suggest that the degree to which anxiety or distrust can be reduced in teacher evaluation should be an important yardstick for measuring an appraisal system's overall effectiveness. The system is more likely to be perceived positively if teachers participate actively in determining the criteria upon which they will be evaluated; teachers are given timely, constructive feedback; there is some system of recourse if a particular appraisal is perceived as unfair; the system has the flexibility to incorporate individual methods and techniques; and teachers are rewarded for meeting the goals or standards of excellent performance. While it may be unrealistic to attempt to completely eliminate the anxiety associated with evaluation, if the system is perceived as fair, valid, relatively objective and systematic then the process will be acceptable to its users.

13.3 Criteria Ambiguity

The evaluation of teaching is constrained by the ambiguity surrounding how to define and measure good teaching. Johnson (1980) points out that "after eighty years of research, there is virtually no consensus among either educational psychologists or practitioners about what constitutes teaching effectiveness."[145] Similarly, Rosenshine and Furst (1973) observe that "a complete list of educational 'shoulds' can only be guessed at ... research in this area has barely begun."[146]

Implications

The search for a good evaluation system cannot be abandoned until a full set of unequivocal criteria can be established. In fact, while research and practice may bring us closer to an understanding of some of the components of effective teaching, there is no reason to expect that the time will ever come when educators possess a comprehensive profile of a good teacher. This is because:

---

[145] Susan Moore Johnson, "Performance-Based Staff Layoffs in the Public Schools: Implementation and Outcomes", Harvard Educational Review, May 1980, p. 230.

[146] Barak Rosenshine and Norma Frust, "The Use of Direct Observation to Study Teaching" in Robert M. W. Travers ed., The Second Handbook of Research on Teaching (Chicago, Ill.: Rand McNally, 1973), p. 163.

> ...the word 'teaching' refers to a broad class of activities. The particular activities that constitute teaching in any particular situation depend upon how the school is organized, the nature of the program, the structure of the curriculum, the teaching materials to be used, the expectations of parents, and the social context of education. A method suitable for evaluating teacher effectiveness in one situation may be quite unsuitable in another.[147]

Teaching criteria should not, and indeed cannot, be highly standardized. What, then, is the alternative?

As subsequent discussion will elaborate, there are several alternative approaches to conducting evaluation.[148] However, a general framework, as illustrated in figure C-2, is suggested here.

Figure C-2

Framework For Establishing Standards

```
                     ______________________________________________
                    |                                              |
Ministry of------|----> Baseline Criteria and Process Guidelines   |
Education           |                                              |
                    |                                              |
                    |        ------------------------------------- |
                    |       |                                   |  |
School Board-------|------|---> General Performance Criteria    |  |
                    |       |                                   |  |
                    |       |   ------------------------------  |  |
                    |       |  |                              | |  |
                    |       |  |                              | |  |
Local School-------|------|->|Specific Performance Objectives | |  |
                    |       |  |                              | |  |
                    |       |  |    ------------------------  | |  |
                    |       |  |   |                        | | |  |
Teacher-----------|------|--|---|---> Mutually-Derived      | | |  |
                    |       |  |   |        Goals           | | |  |
                    |       |  |   |                        | | |  |
                    |_______|__|___|________________________|_|_|__|
```

It may be feasible for the central administrative office to establish province-wide baseline criteria for teacher effectiveness or standardized guidelines for the process of appraisal to ensure consistent implementation across the province (e.g. what general approach to use, how often to evaluate etc.). School boards, working within these general guidelines,

---

147. Robert M. W. Travers, "Criteria of Good Teaching", in Jason Millman ed., Handbook of Teacher Evaluation (Beverly Hills, Calif.: Sage Publications, 1981), p.22.

148. See section 14.0.

could develop more specific criteria which embody the needs and expectations of their own unique constituencies. These criteria would, in turn, still allow sufficient latitude at the local school level to set performance objectives on the basis of those criteria which are responsive to the unique needs and concerns of the particular school and its staff. Supervisors and teachers, working together, could develop mutually-derived goals. In this way, working within a general framework, evaluation at the teaching level would be linked with curriculum and program objectives which are responsive to the local community and would still allow for flexibility to meet unique individual and institutional requirements.

13.4 Teaching as a Contextual Experience

The teacher operates within an environment that strongly influences his or her level of performance. This environment includes the economic and social conditions of the community, community expectations, the kinds of students enrolled in the school, the financial support, institutional arrangements, facilities and atmosphere of the school, other staff, and classroom situations. Performance is hindered or enhanced by external factors relating to the physical classroom setting, availability of equipment, the size, capacity, interest, abilities and work habits of the class, quality of materials of instruction, and the curriculum. All these factors confound the process of evaluation.

Implications

As part of the supervisory process, the evaluation of performance should embody not only the assessment of behaviour but also the means for improvement. The context within which the teacher functions must be acknowledged in the evaluation process and arranged or altered by the supervisor to remove obstacles to ineffective performance and provide opportunities for growth and performance improvement. The teacher does not operate in a vacuum. His or her potential for improvement is largely determined by or dependent upon the teaching environment; it is the supervisor's responsibility to recognize this in the performance appraisal and to change those elements of the environment within his or her control in order to aid teaching effectiveness.

13.5 The Teacher as Professional

The National School Public Relations Association (U.S.A.) claims that the core problem of evaluating teacher performance is the difficulty of making valid judgements about anything as complex and personal as teaching ability.[149] The role of the teacher is characterized by autonomy, isolation, initiative, and task complexity. Therefore, the professional nature of teaching itself constrains the process of appraisal because work cannot be continually observed, teachers exercise a great deal of discretion in their jobs, and they have considerable latitude in developing highly personalized techniques and styles of teaching.

Implications

Many evaluation programs fail in the educational setting because attempts are made to regulate and systematize those elements of the job which defy standardization. As indicated earlier, it is more reasonable to expect a teacher appraisal program to take account of individual, organizational, and community differences, and to standardize only those criteria which pertain to teaching performance independent of context or individual.

Glasman (1977) has pointed out that:

> ... decisions which are based on results of performance evaluation and which entail some kind of a comparison or a differentiation of individuals in terms of their performance are possible only if the individuals who are evaluated are engaged in pursuing similar tasks, and if these tasks constitute similar proportions of their total obligations to the institution.[150]

Since the professional nature of teaching necessitates individualized styles and techniques, the ability to draw explicit comparisons between teachers for purposes of summative evaluation may be difficult. This is one reason that at the classroom level assessment measured against mutually-derived goals rather than solely against a set of prescribed criteria is frequently advocated as a more realistic method of appraisal.

---

[149] Educational Research Service, Evaluating Teacher Performance, p.7

[150] Naftaly S. Glasman, "A Proposed Structure For Evaluating Personnel For Decision-Making", Studies in Educational Evaluation, Spring 1977, p. 49.

13.6 Evaluator Resistance and Ability

Constraints on the process of evaluation can stem from an evaluator's unwillingness or inability to assess performance. With regard to evaluation by principals, for example, Johnson (1980) states:

> There is something in the nature of teaching and evaluation, in the roles of principal and teacher, and in the traditional organization of the school that accounts for the reluctance, ambivalence and resistance of building principals to embrace the task of competitively assessing teachers' work - a responsibility that is thought to be routinely accepted by supervisors in business.[151]

Wolcott (1973), in his ethnographic study of the role of the principal, offers one explanation for the resistance:

> ... the act of evaluation conflicted blatantly with the ideal of democratic administration in which most principals preferred to present themselves among their teacher colleagues as a first among equals.[152]

Other writers have suggested more specific reasons for evaluator resistance:

* A general lack of certainty regarding criteria and procedures;
* the evaluator's dislike of being in a position to manipulate or adversely affect other people's lives;
* a fear of precipitating an unpleasant reaction on the part of the person being evaluated;
* a failure to see the evaluation of others as relevant to the purposes of the evaluator; and
* an inability to organize time so that adequate assessments can be made.[153]

Nor is evaluator resistance the only limiting factor; many of the people responsible for teacher evaluation lack technical knowledge or expertise regarding measurement and evaluation. Castetter (1976) believes that in order to perform the evaluation task effectively, administrators must possess a variety of human, conceptual, and technical skills which, for the most part, must be developed.[154]

---

151. Johnson, "Performance-Based Staff Layoffs", p.224.

152. Harry F. Wolcott, The Man in the Principal's Office: An Ethnography (New York: Holt Rinehart, 1973), p. 194.

153. See Goodwin, "Constraints on Teacher Education", p.27.

154. W. B. Castetter, The Personnel Function in Educational Administration (New York: MacMillan, 1976), p. 246.

Implications

It is evident that evaluator training could do much to ameliorate this problem. This may serve to reduce both the unwillingness and inability to conduct appraisals. Nevertheless, a good performance appraisal instrument should have the effect of minimizing evaluator discomfort with the process and maximizing the ability of principals to make fair and accurate assessments of performance. Ertis (1980) states that "the problem of variation in the standards of different evaluators has never been resolved ...";[155] yet the purpose of an appraisal system is not to eliminate the standards of different evaluators, but rather to extricate those standards from the minds of evaluators (where bias and subjectivity are a constant threat) and place them within a systematically derived appraisal instrument, which different evaluators can use to simplify, objectify, and legitimize the process of assessment.

14.0 Approaches to Teacher Evaluation

14.1 Current Conventions

In the vast majority of schools in Canada and the United States, observation of the teacher by the principal and/or supervisor, using some form of checklist or rating scale, is the standard method of evaluation. Sometimes the evaluator informs the teacher ahead of time that he or she will be observing the class, and may or may not hold a formal pre-observation conference with the teacher to discuss the objectives and process of the evaluation. During observation, the evaluator takes notes on what is seen and heard, and ordinarily has a rating form to assist in identifying characteristics or behaviours. Subsequent to the observation, a post-observation conference is usually conducted, in which the evaluator and the teacher discuss the results of the evaluation. The rating form may or may not be signed, and may or may not be consulted or circulated again once it has been placed in the individual's personnel file.

---

[155] Bernie P. A. Ertis, "How Much Are You Really Worth?", OSSTF Forum, December 1980, p. 203.

In recent years, this approach has been subject to increasing disapprobation on the basis of questionable validity and lack of acceptability to both the evaluator and the evaluatee.

> Observations may not be good measures of teaching effectiveness because the results aren't generalizable beyond the situation under observation, because they are subject to observer bias, because they are expensive in time and money, and because atypical behaviour of the teacher is often demonstrated when the observer is present.[156]

In an effort to improve the validity and reliability of the classroom observation approach, a number of process assessment instruments, such as the Flanders Interaction Analysis System, the Apple System, and the Ramos system have been developed to assist the evaluator in collecting objective data about categories of behaviour in classrooms.[157] Generally speaking, such observation instruments have been, in comparison with current rating systems, carefully developed and validated. Evidence indicates that observation instruments are very useful for feedback purposes. However, the use of such instruments for evaluative decision-making is presently quite rare.

With regard to the observation approach in general, Frederick McDonald (1980) provides a useful list of principles, rules and errors to be avoided in classroom observation.[158] These are reproduced in exhibit C-2.

---

156. Robin H. Farquhar, "How the Teaching Profession Measures Teaching Effectiveness". (Paper presented at the Conference on Teaching in Medical School, Saskatchewan 1978), p. 9.

157. See exhibit C-3 for examples of these systems.

158. Frederick McDonald. "Principles and Procedures in Observing Classroom Instinction" in Donovan Peterson and Annie Ward, eds., Due Process in Teacher Evaluation (Washington, D.C.: University Press of America Inc., 1980), pp.110-113.

Exhibit C-2

Principles, Rules and Errors to be Avoided in Classroom Observation

The points in this summary are guidelines for the development and operation of a system for evaluating the classroom instructions of teachers. They are separated into three groups. The first group is a set of principles that should be used in designing a system for observing teachers. The second group is a list of rules, stated rather directively; these rules, if used, ought to help you function effectively as an evaluator. The third group is a list of errors to avoid.

Principles to be Used in Developing an Observation System

1. The goal of the observational system is to establish the characteristic level of performance, the characteristic style of a teacher.

2. The observation is a sample of teaching. As such, it must be a representative sample of the teacher's characteristic way of teaching.

3. The observational data will vary depending on who is observing and the characteristics of the occasion of observation -- time of day, day of the week, what is being taught. A sufficient number of observations must be made to allow for these variations.

4. A teacher must be observed frequently and under different circumstances to describe the teacher's characteristic style.

5. The accuracy of the observations should be established by having at least two persons observing independently and simultaneously on several occasions.

6. The substance of what is observed should be those teaching performances which are known to be directly related to pupils' learning and growth.

7. The results of the observations should have "face validity" for the teacher; i.e., the teacher ought to be able to recognize her/his teaching style in the observational record.

Some Rules to be Used in Observing Teachers

1. As the first step in developing an observation system, describe those teaching performances directly related to pupil learning. These descriptions should be developed with teachers and agreed to by them.

2. The teaching performances or activities should be those which are known to be or for which there is good reason to believe are necessary to achieve the major goals in instruction.

3. The list or description of desirable teaching performances and activities should cover: (1) how instruction is organized -- the size and composition of teaching groups; (2) the quality of instruction --the clarity of directions, explanations, questioning strategies, etc., (3) opportunity to learn -- variety and quality of learning experiences, content covered, opportunity to learn skills useful or necessary for further learning, etc.; (4) evaluation of pupil learning; (5) adaptation of learning experiences and materials to the aptitudes and learning styles of pupils; and (6) management of pupil involvement in instruction.

4. Categories cited in No.3 (above) include many different kinds of specific teaching performances. If a practice is directly related to the goal of the category, for example, increasing the amount of direct instruction or providing more constructive evaluation, it should be regarded as a required practice. There probably will be different ways of exemplifying in action the functions which define these categories. Some may be more effective than others in achieving the goal of a function. The data should tell the evaluator whether or not the teacher is carrying out the critical functions of instruction.

5. Only those teaching performances should be considered which the teacher can control. These performances are: how instruction is organized (large group, small groups, whole class), interactions with pupils, explanations, directions, supervision of seatwork, accessibility of teacher to give help, control of ontask behaviour, evaluation of work.

6. Note for each observation the activity of pupils with whom the teacher is not working directly, are they productively engaged? Treat such observations as only potentially related to ineffectiveness practice.

7. Organize the observation schedule in terms of significant periods of extended teaching -- a sequence in reading or mathematics, a special unit.

8. Arrange these periods of observation throughout the year, varying where possible, the day of the week and the time of the day.

9. In this set of observations, provide for at least two observers, for example, the principal and the assistant principal. At least two of the observations should be made simultaneously by both observers.

10. The number of observations should not be less than five.

11. Observations should be discussed with the teacher. These discussions are necessary to determine what the teacher was trying to accomplish and why she/he did certain things.

12. Observations made early in the year should be used for the teacher's development.

13. The criterion of evaluation should be whether or not or to what degree the teacher uses an effective style.

14. An equally important criterion should be the extent to which the teacher improves over time and as a result of the discussion of the early observations.

15. The goal of the observations should be to establish the teacher's characteristic style. The question to be answered is, does the teacher have a teaching style likely to be effective in stimulating pupil learning and growth? Counting frequencies of a few teaching behaviours is not likely to assess the teacher's style. The observation data should tell you what the teacher does to achieve certain goals under specific conditions and how she/he copes with instructional problems.

Errors to be Avoided

1. A few observations are probably useless.

2. The observational data will not be interpretable if the teacher's goals and the classroom setting are not known.

3. The data should be interpreted with the teacher. Failure to do so means that the decisions the teacher has made while teaching are unknown. Her/His problem may be in how she/he thinks about teaching.

4. Do not set the evaluation standard so high that only the exceptional teacher can meet it.

Reproduced from: Frederick McDonald, "Principles and Procedures in Observing Classroom Instruction" in Donovan Peterson, and Annie Ward eds., Due Process in Teacher Evaluation (Washington, D.C.: University Press of America Inc., 1980), pp. 110-113.

Classroom observation is currently the most typical and widespread approach to teacher evaluation. Although adherence to consistent, reliable methods of observation and selection of suitable rating scales contribute to the validity and acceptability of such an approach,[159] McDonald and others are now suggesting that classroom observation is only one potential component in a system for gathering data about teaching effectiveness. As Evertson and Holley (1981) state, "classroom observation is only one, albeit a very important, source of relevant information for teacher evaluation and must be supplemented by other sources of information...."[160]

Therefore, shortcomings in this approach which cannot be ameliorated by better rating scales and observational techniques may be handled either by replacement of this approach with another or by the addition of other techniques or appraisal sources to offset inherent deficiencies of observation.

14.2 Variations to Conventional Approaches

Variations to the conventional approach can occur along one of three dimensions: sources of appraisal, choice of criteria, and type of technique. Sources of appraisal refer to which individuals will have input into the evaluation process or serve as a source of data on teacher effectiveness. Choice of criteria involves decisions to use process or output criteria (teacher behaviour or student learnings) and performance, skills or trait characteristics of effective teaching. The third dimension refers to the specific technique used in the appraisal process. In section 7 of this report, these various techniques were described in general. In the area of teaching, these techniques commonly fall into one of four categories: specific instruments, rating scales, checklists, and narratives.

Although specific "packaged" instruments are not widely used, the purported merits and difficulties associated with them have been given fairly extensive study. The most commonly used instruments are outlined in exhibit C-3, along with a summarized description of their context and a suggested source reference for the reader who is interested in learning more about a particular technique. The overriding contribution of these techniques is a result of their orientation towards systematization of procedures to minimize the influence of evaluator bias. At the same time, as noted earlier, most instruments are useful for developmental and feedback purposes, rather than summative, ones. In addition, they usually require skills based on special training (which may be costly or time-consuming). Finally, there is always the danger with these techniques that they will "take on a life of their own" and either cloud the overall purposes of evaluation or be applied to inappropriate settings and circumstances.

---

[159] Charles Mitsakos, "Generating a Valid and Reliable Rating Scale For the Assessment of Teaching Performance", in Robert J. Munnelly, ed., The Improvement of Local Staff Evaluation and Supervision, (Massachusetts Association For Supervision and Curriculum Development, May 1981). (ERIC Document Reproduction Service No. ED 208 525).

[160] Carolyn M. Evertson and Freda M. Holley, "Classroom Observation" in Jason Millman, [ed.], Handbook of Teacher Evaluation (Beverly Hills, California: Sage Publications, 1981), p. 94.

Exhibit C-3

Common Classroom Assessment Instruments: A Summary

INSTRUMENT	CONTENTS/PROCESS	SUGGESTED SOURCE REFERENCE
Flanders Interactional Analysis	- analyzes verbal interaction between teachers and students	Flanders, Ned A. Analyzing Teaching Behavior. Reading, Massachusetts: Addison- Wesley, 1970.
Galloway Non-Verbal Communication	- analyzes the types of non-verbal behaviour teachers use with students by means of video tape recording	Galloway, Charles M. "The Nonverbal Realities of Classroom Life in Observational in the Classroom, C.W. Beegle and R.M. Brandt [eds.] Washington, D.C.: Association For Supervision and Curriculum Development, 1973.
Parson's Types of Analysis	- analyzes the types of questions teachers use with students	Evaluation Handbook. Fairfax, Virginia: Fairfax County Public Schools, 1975.
Bales Interaction Process Analysis	- analyzes interaction patterns of group members	Evaluation Handbook (see above)
Verbal Interaction Category System	- an adaptation of the Flanders system, including measurement of non-verbal behaviour of teacher and student	Griffith, Francis. A Handbook for the Observation of Teaching and Learning. Midland, Michigan: Pendell Publishing Company, 1973.
Classroom Observation Record	- analyzes cognitive levels on which classroom verbal interaction takes place	Griffith, Francis. (see above)
Observation Guides	- comprehensive itemization of specific and observable aspects of teaching and learning which helps supervisor to monitor certain phases of instruction	Griffith, Francis. (see above)
Briggs Observation Guide	- a collection of questions which serves as a guide to help supervisors arrive at judgments regarding the purpose of a lesson, classroom climate, organization and development of lessons.	Griffith, Francis. (see above)
Instrument for Observation of Teaching Activities (IOTA)	- a written description of classroom behavior by a team of at least three observers	"Observation Procedures" in Staff Evaluation Skills. Academy Skill Institute 75-S-2. Arlington, Virginia: American Association of School Administrators, January 1975.

Video-tape Analysis	- a system which allows teacher and supervisor to review a lesson and reach consensus on constructive alternatives for teaching improvement	Griffith, Francis. (see above)
Observation Schedule and Record System (OSCAR)	- a verbal category system which yields frequency counts of the occurrence of of different verbal behaviours.	Medley, Donald M. "Measuring the Complex Classroom of Today" in Observational Methods in the Classroom. (see above).

Derived from: Educational Research Service, Evaluating Teacher Performance, (Arlington, Virginia: ERS Inc., 1978).

There are a wide variety of rating scales in use today. All scales share certain features in common. Generally they contain a description of effective teaching (which may be based on traits, characteristics, skills, competencies or behaviours). Using this description, the rater determines to what degree a teacher manifests the described quality or behaviour. The degree or extent of manifestation is scaled along a continuum which usually incorporates a number scale associated with a verbal indicator (poor, needs improvement, satisfactory, superior, outstanding, for example). The rater judges the teacher by putting a check on the number scale or circling the most appropriate description. Exhibit C-4 reproduces an example of a rating scale used in classroom observation. Many educators question the validity of the rating scale.[161] Section 7.2.2 of this report has outlined the advantages and disadvantages of this technique, so they will not be repeated here.

The checklist, like the rating scale, is comprised of items which describe various aspects of effective teaching. The second part of exhibit C-4 illustrates this technique. This type of judgement is qualitative in that the existence or absence of a particular attribute is acknowledged without reference to the degree to which it prevails.

---

[161] See, for example, Richard M. Brandt, "Toward a Taxonomy of Observational Information", in Charles W. Beegle and Richard M. Brant eds., Observational Methods in the Classroom (Washington, D.C.: Association For Supervision and Curriculum Development, 1973), pp.23-34.; or James Popham, "Pitfalls and Pratfalls of Teacher Evaluation", Educational Leadership, November 1974, pp. 141-146.

Exhibit C-4

Sample Rating Scale

OBSERVATION GUIDE
WEST DES MOINES TEACHER PERFORMANCE

(Teacher's Name) ______________________ (Date) ______________

(Class Observed) ______________________

OBSERVER/EVALUATOR: ______________________

BRIEF DESCRIPTION OF ACTIVITY OBSERVED

	DEGREE OF COMPETENCY EXHIBITED*				
Check one or more comments for each item.	Optimum	Average	Not shown	Seek more information	Notes
(1) Teaching toward perceivable objectives. Teaching focused on a particular learning target.					
(2) Objective(s) appropriate for learners — not already attained or too difficult.					
(3) Was progress toward objective(s) made? If objective(s) switched in period, was change for valid reasons?					

Exhibit C-4

Sample Rating Scale

	DEGREE OF COMPETENCY EXHIBITED				
	Optimum	Average	Not shown	Seek more information	Notes
(4) Teacher's behavior facilitated learning.					
(5) Teacher minimized interference with learning.					
(6) Supporting materials and media well prepared/chosen.					
(7) Evaluation activities effective.					
(8) Classroom environment created which is conducive to learning.					
(9) Provides a variety of learning modes.					
(10) Appropriate applications of principles of learning.					

(11) What were the best performance areas?

(Check one or more)

( ) Course content selected
( ) Objectives
( ) Methods
( ) Student-student interaction
( ) Student-teacher interaction
( ) Student independent study
( ) Other: ____________________

(12) What were the performance areas needing improvement?

(Check one or more)

( ) Course content selected
( ) Objectives
( ) Methods
( ) Student-student interaction
( ) Student-teacher interaction
( ) Student independent study
( ) Other: ____________________

Reproduced from: Beatrice M. Gudridge, Teacher Competency: Problems and Solutions (Sacramento, California: American Association of School Administrators, Education News Services, 1980), pp. 43-44.

Griffith (1973) suggests several advantages to this approach:

1. It directs attention to aspects of a lesson which an observer might otherwise miss.
2. It gives a degree of objectivity to an evaluator's observations.
3. It provides a permanent record which is quick and easy to make.
4. It helps a teacher to analyze his or her own lesson and to determine what a supervisor considers important.[162]

There are, however, a number of shortcomings to the checklist:

1. A checklist influences an evaluator to analyze teacher performance during a lesson according to a common pattern even though lessons may vary widely in form and purpose. As a result, classroom observation may become a mechanical, routine procedure.
2. Items on a checklist often are numerous and vary in significance. There is rarely any attempt to weigh their relative importance.
3. Checklists usually deal with details which are often superficial.
4. When the use of checklists becomes routine, supervisors are apt to make judgements without patient reflection and careful analysis.[163]

The narrative technique corresponds to "essay evaluations" outlined in section 7.2.1. This may be a written report of the teacher's strengths, weaknesses, behaviours, competencies, areas in need of improvement, or future potential. The rich descriptive nature of this technique and its use for feedback purposes make it a worthy supplement to other types of techniques. However, as noted earlier, this technique is of limited usefulness for making evaluative decisions.

The selection of appropriate appraiser(s), criteria, and techniques all have significant impact on the quality of classroom observation methods and may in some instances provide useful supplemental data for use in the evaluation process. Nevertheless, alternative approaches to teacher evaluation have been proposed and in some instances implemented with relative success. It is these approaches to which we now turn.

---

162. Francis Griffith, A Handbook For the Observation of Teaching and Learning, (Midland, Michigan: Pendell Publishing Company 1973), p. 54.

163. Ibid.

14.3 Innovations in Teacher Evaluation

Three approaches which diverge radically from conventional methods are: teacher testing, the performance-objectives approach, and the use of behaviourally anchored rating scales.

14.3.1 Teacher testing

The testing of teachers has recently moved within the domain of performance appraisal. Traditionally, administering a test to teachers has been considered within the confines of screening job applicants and making selection decisions; but as recently as 1979 the teacher testing movement has gained sufficient momentum in the United States. Advocates suggest its use on veteran teachers to determine their competency. The National Teacher Examinations (NTE) have been the most widely used tests for new teachers in the United States and five states now use the NTE for certification purposes.[164] In a U.S. Supreme Court case, the court ruled that NTE scores are related to teacher's job performance. A content validation study conducted by the Educational Testing Service met the content requirement for a validated test. Nevertheless, considerable controversy surrounds the use of tests for performance appraisal purposes. As Haefele (1980) concludes:

> The NTE is an achievement test. As the Coleman report and other studies indicate, variation in levels of teacher knowledge of content in education is not clearly associated with successful teaching. Until research appears showing stronger relationships with practice, the practical utility of the National Teacher Examination is questionable.[165]

A few authors (Popham, 1971, for example) have also advocated the development of in-school "teaching tests" as a method of teacher evaluation.[166] As Levin (1979) points out, Popham suggests that "a number of teachers be given one or more identical objectives and samples of the measures that will be used to assess them. Teachers would then plan and teach a lesson to a controlled group of students under controlled conditions."[167] After instruction, "measurement tools are administered and the students' scores on the posttest become an index of the teachers' effectiveness."[168] This technique, although intuitively attractive, contains several problems:

---

[164] Louisiana, Mississippi, North Carolina, South Carolina and West Virginia.

[165] Haefele, "How To Evaluate Thee", p.351.

[166] James Popham, "Performance Tests of Teacher Proficiency", American Educational Research Journal, January 1971, pp.105-17.

[167] Levin, "Teacher Evaluation", p.243.

- The objective may not be related to the subject area of the teachers;
- Random assignment of students that the teacher knows little or nothing about could initially affect instructional effectiveness;
- We must assume that these students reflect levels of motivation, judgement etc. typical of students the teacher interacts with in class; and
- The ability level of the assigned group could make a difference since some teachers are particularly effective with high-ability as opposed to low-ability students, or vice versa.[169]

In addition, as Glass (1974) argues, the approach has to be able to demonstrate high reliability across different situational topics and different pupils.[170]

14.3.2 The Performance Objectives Approach

In section 12.3 we introduced the notion of goal setting as a potentially valuable component of evaluation. Based on this principle, George Redfern (1980) has developed and refined the "Performance Objectives Approach" to the evaluation of educational personnel.[171] Since 1975, Redfern has assisted at least sixteen American school districts in implementing this approach to evaluation.

The approach involves a dialogue between a supervisor and teacher in which the goals and objectives of the teacher for the approaching period of time (semester or academic year, for example) are mutually derived.[172] Activities involved in achieving these goals are defined and an evaluation is made based on the extent to which the teacher has achieved the objectives agreed upon in the previous goal-setting interview. Redfern outlines a six stage evaluation model (see exhibit C-5). A brief description of each of the components is also presented.

---

[168] Ibid.

[169] Haefele, "How To Evaluate Thee", p.251.

[170] Gene Glass, "Teacher Effectiveness", in Herbert Walberg ed., Evaluating Educational Performance (Berkeley, California: McCutchan, 1974), p. 11-32.

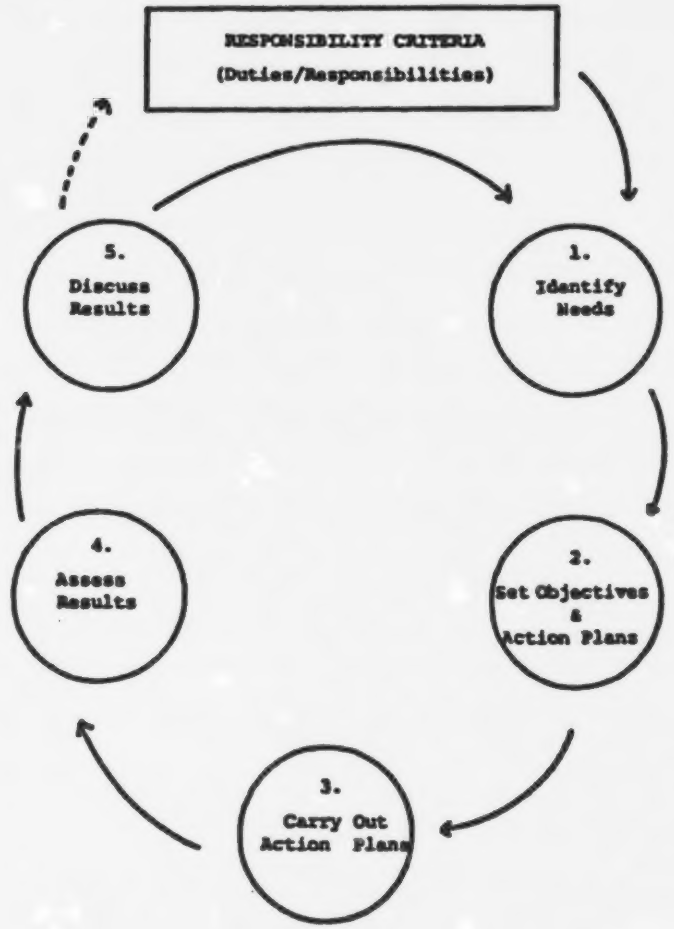
[171] This approach is not limited to teacher evaluation only. See George B. Redfern, Evaluating Teachers and Administrators: A Performance Objectives Approach (Boulder, Colorado: Westview press, 1980).

[172] Note that this approach closely approximates management by objectives (MBO) as discussed in section 7.4.2.

Exhibit C-5

Performance Objectives Approach

Evaluation Model


RESPONSIBILITY CRITERIA
(Duties/Responsibilities)
1. Identify Needs
2. Set Objectives & Action Plans
3. Carry Out Action Plans
4. Assess Results
5. Discuss Results

Reproduced from: George B. Redfern, Evaluating Teachers and Administrators: A Performance Objectives Approach (Boulder, Colorado: Westview Press, 1980), p.14.

(1) Responsibility Criteria

Duties and responsibilities required in the performance of an assignment must be indicated. Most school systems have developed criteria of one sort or another that can be used for evaluating performance.

(2) Identify Needs

Using responsibility criteria, the evaluatee and the evaluator cooperatively identify the status of the former's current performance. The identification of current strengths and areas of difficulty can be useful in deciding the kinds and number of specific objectives and action plans to establish.

(3) Set Objectives and Action Plans

Objectives and action plans are the means to achieve desired outcomes determined by the evaluation process. Several questions arise as objectives are being formulated. How will they be stated? In behavioural terms? Or in terms of learner achievement? Or both? Will the attempt be to effect a change of behavior on the part of the teacher or the pupil? Or both? If learner achievement is the consideration, will the performance objectives be directed toward gains in both the cognitive and the affective domains? (The former are easier to design and to measure than the latter.) The art of drafting objectives must be cultivated.

At the time performance objectives are agreed upon, it is important to discuss the actions and efforts that will be expended to attain the objectives. If the objectives concentrate upon changes in the behavior of the evaluatee, what types of changes will these be? Alterations in teaching or administrative style and tactics? Better planning? Greater flexibility and adjustability? Increased emphasis upon innovation and creativity? More realistic evaluation of pupil achievement? If these are to be objectives, both the evaluatee and the evaluator must plan actions and activities calculated to bring about the desired changes in teacher behavior.

If the performance objectives include achievement gains of the learner, the emphasis will be upon actions to bring about such gains. The evaluatee and evaluator will decide what instructional and administrative procedures will be employed. There will be a concentration upon a better grasp of concepts and an emphasis upon skill development. The evaluatee and the evaluator pursue these actions cooperatively. Both have a vested interest in the outcome.

(4) Carry Out Action Plans

Traditionally, monitoring and evaluating have been done simultaneously, and observation of teaching performance has been essentially an evaluative process. Little, if any, effort has been made to distinguish between monitoring and evaluation. It has been like deciding the winner of a game long before it has ended.

The evaluator should monitor the evaluatee's performance to collect data and information that relate to the objectives being pursued. Monitoring is concerned with performance outputs; it is the evidence-gathering part of the total evaluation process. The parties involved must discuss and, it is to be hoped, agree upon certain matters concerning the monitoring (i.e., the data-gathering forms to be used, kinds and frequency of visitations, identity of monitors, and the use of any mechanical monitors, conferences, and other types of contacts.). A substantial volume of appropriate information enhances good evaluation and may also result in immediate change for improvement. Information from the monitoring should never be simply stored away when prompt feedback will enhance performance.

(5) Assess Results

Interpreting the meaning and significance of monitored data is a very important part of the total process of evaluation. This represents the culmination of all that has gone before. Goal setting, the determination of objectives, and performance activities all lead up to the evaluation stage.

Evaluation of data is a twofold process. Self-assessment is one part; the evaluator's assessments compose the other. If more than one evaluator is involved, assessment can become a three-part or even a collegial process. Some students of the teacher evaluation process are coming to believe that assessment can and should be a collegial process. Most performance requires primary and secondary evaluators, but various forms of collegial assessment may be used on an experimental basis.

The classic one-on-one, evaluator-evaluatee type of assessment is not very realistic in a large school where there may be dozens of teachers, one principal, and little other administrative or supervisory staff. Also, if the cycle of evaluation periods must begin and end for all teachers at the same time, utilizing a single evaluator is unrealistic when the evaluator-evaluatee ratio is high.

What are the options in the involvement of others in the assessment of monitored data? One possibility is to use a supervisor or consultant to aid the principal in a review of the monitored data. In some secondary schools, the department heads also assess monitored data. Another pattern that has been suggested would involve creating a committee of teachers to assist in the review.

There is mounting interest in having students contribute to the assessment process. In some instances, this has taken the form of student demands; in others, educators have resolved to find better ways to involve the students. Also, in some school systems there are demands that parents should have some role in evaluating the services of teachers and administrators. Obviously the demands for broad participation make a performance-oriented evaluation program essential if the abuses and errors of mass involvement in subjective determinations are to be avoided.

(6) Discuss Results

The evaluation conference is exceedingly important. It is the occasion for the persons most intimately involved in the process to discuss the outcome of their efforts to achieve the objectives. A very important responsibility is placed upon the evaluator to help the evaluatee view evaluation as a constructive rather than a negative process.

There is no set recipe for conducting a productive evaluation conference. Preparation will help. Training is obligatory. If both parties have tried to the fullest to achieve the performance objectives, that fact will contribute to the success of the conference. The most valuable asset, however, is experience on the part of the evaluator.

The evaluation conference will yield ideas for follow-up action. It may develop, during the conference, that the evaluatee will see the need for certain kinds of follow-up activities to reinforce gains already made. The tentative formulation of objectives for the next round of evaluation often takes place during the conference.

> Performance objectives evaluation can be the means of enabling individuals to achieve goals that will contribute to improved performance. It is a tool to help people become more competent in the performance of their duties and responsibilities. Those duties and responsibilities must continually be evaluated in relationship to the primary task of any school -- that of improving learning opportunities for the students. Evaluation is a worth-while enterprise and deserves more penetrating and continued study.[173]

While it may be too early to fully assess the feasibility of the goal-setting approach, it is evident that it provides a comprehensive alternative to traditional methods of teacher evaluation. The potency of goal setting as a motivational tool suggests that this approach may, in fact, go as far (or further) than other methods in establishing procedures to improve performance and enhance professional development. It may become apparent over time, and with repeated replication in different schools, that the applicability of such an approach is contingent upon the chosen philosophies, styles, and objectives of particular schools. Hence, while it may lack universal application, there may be certain schools for which, in comparison with other methods, the goal-setting approach is optimally suited.

14.3.3 Behaviourally Anchored Rating Scales

A final innovative approach to evaluation bears mentioning because it exemplifies a serious attempt to objectify teacher assessment on the basis of behavioural characteristics. This is the use of Behaviourally Anchored Rating Scales (BARS) in teacher evaluation. This technique was reviewed in section 7.4 of this report, so it remains here only to provide an example applicable to teachers. Appendix 1 illustrates behaviourally anchored rating scales for each of the following eight dimensions of evaluation for special education teachers:

1. Professional Growth
2. Communication Skills
3. Rapport with Students
4. Appropriate Planning
5. Instructional Skills
6. Record Keeping and Reporting

---

[173] Redfern, Evaluating Teachers, p. 13.

7. Behaviour Management
8. Interstaff Relationships.[174]

The use of such a technique has not been generally recognized in the area of education. As in industry, this may be due in part to the amount of time, cost and effort which must be allocated to its design. While its application in the private sector has met with mixed results, it indicates the direction in which several performance appraisal methods are moving. That is, the growing recognition (through both research and practice) that traits or personality criteria lead to inadequate and biased assessment has impelled many evaluation system designers to identify standards which relate to specific, observable competencies and behaviours of effective teaching. The BARS approach illustrates a systematic and comprehensive effort to achieve just this goal.

15.0 Summary

Eight reasons as to why a teacher evaluation system is needed have been outlined. These reflect the intention of appraisal, which is to serve the requirements of all upon whom the teaching process impacts: students, teachers, administrators, parents, and the general public. Critical issues surrounding teacher evaluation relate to the question of whether to use presage, process or output criteria; the choice of who should conduct the appraisal; the role of goal setting in evaluation; and the issue of formative versus summative evaluation. In order to guide the design and use of an evaluation system, twenty-five components of a good evaluation system were summarized. These, however, may be subject to various constraints on the process which tend to characterize teacher appraisal, specifically: lack of validated research, teacher resistance, criteria ambiguity, the contextual experience of teaching, the teacher as professional, and evaluator resistance and ability.

A number of approaches to teacher evaluation have been summarized. The current convention is classroom observation by the principal and/or supervisor using some form of rating scale. Variations to this approach occur along one or more of three dimensions: sources of appraisal, choice of criteria, and type of technique. Four techniques were identified: specific instruments, rating scales, checklists, and narratives. Finally, three innovative approaches to teacher evaluation were described: teacher testing, the performance objectives approach, and behaviourally anchored rating scales.

---

[174] From Robert J. Beebe, "Use of the Behaviourally Anchored Rating Scale in Evaluating Teacher Performance". (Paper presented at the Annual Conference of the Evaluation Network, Tennessee, 1980).

It is evident that while strides are being made in the design, use and implementation of teacher performance appraisal, there remains a considerable degree of controversy surrounding the subject.

> When one finds ... that competent observers advocate strongly divergent points of view, it seems likely on a priori grounds that both have observed something valid about the natural situation, and that both represent a part of the truth. The stronger the controversy, the more likely this is.[175]

The controversial nature of teacher evaluation signifies not only its complexity but also the improbability of designing highly prescribed and detailed vehicles for the assessment of teaching effectiveness on a universal basis. Like the process of teaching itself, appraisal methods must of necessity be somewhat "individualized" and "stylized" to fit the needs of the particular classroom, school, and local constituency. It is in this way that different methods of teacher evaluation each "represent a part of the truth".

This is not to suggest that universal standards of teacher effectiveness upon which there is consensus are without practical purpose. Such standards serve as important guidelines in teaching performance and clarify to both educators and the public alike the parameters of accountability. Rather, these standards should serve as boundaries within which more specific and locally relevant criteria are applied in response to the needs and demands of particular individuals, curricula and school settings. For this reason, a system of appraisal, where different methods and criteria are used both laterally and vertically (i.e., across different districts and at different hierarchical levels) is likely to be more acceptable and more valid than a single, detailed appraisal system applied uniformly to all levels and settings.

---

175. D. T. Campbell and J. C. Stanley, Experimental and Quasi-Experimental Designs For Research (Chicago, Illinois: Rand-McNally, 1963), p. 3.

16.0 Conclusion

Both section B and section C of this report have been reviews of literature first of performance appraisal in general and second on teacher evaluation. Both sets of literature are very similar in identifying several demands or requirements which must be satisfied by an appraisal system if it is to be effectively utilized in an organization. The list of possible uses of an appraisal system is essentially the same for both a business and education setting.

a) The appraisal system can be used to generate data for personnel administration (compensation, training, counselling, and staff planning) and for personnel management decisions (retention, promotion, transfer, discipline). Decisions for personnel administration require appraisal data which indicate whether a person's level of performance is satisfactory to retain their current job, i.e., performance at least satisfies the minimum standards. Decisions for personnel management require data which distinguish levels of performance from one person to another, so that equitable and valid standards are used for decisions on retention, promotion, and discipline.

b) The appraisal system can be used for supervision and motivation of personnel (probationary review, communication, negotiation and clarification of assignments and responsibilities, goal setting, delegation, support and correction). Supervision of personnel via the appraisal system is accomplished through communication and formal discussion of work assignments in a postappraisal interview and informally throughout the year. If a subordinate knows that these assignments will be subsequently reviewed in the next appraisal session, the motivation to concentrate on these assignments should be enhanced. In addition, if the appraisal system provides an opportunity to participate in the setting of assignments so that individuals are working essentially toward their own goals, then motivation should be enhanced as well.

c) The appraisal system can be used to introduce, maintain and/or change operating methods, to clarify, monitor, and identify training needs. If the organization is convinced that a particular philosophy or particular methodology of professional service is superior or desirable, then the appraisal system performance criteria can be designed to elicit and reward this particular behaviour. For example, if an education institution believed that highly theoretical academic research was the path to organization excellence, then the appraisal system would include teachers' publications in referred journals as a performance criterion. If the institution believed that teaching quality was the path to excellence, then student ratings of teachers and courses would be a criterion. If an organization wished to change its philosophy or methodology of professional service, then it would be necessary to develop a new set of performance criteria to elicit and reward the new type of professional behaviour.

d) The appraisal system can be used for the achievement of organizational and personal goals by translating organizational goals and plans into a series of individual goals and plans at all levels of the hierarchy. This approach to appraisal is commonly used where an organization has developed a mission or strategic plan that requires coordinated effort toward its achievement. Departmental and individual goals can be made up of subsets of organization goals. Individual task assignments can be derived from the larger goals. In either case, if individuals are held accountable for performance which contributes to the achievement of an organization's goals, then the likelihood of these goals being achieved is enhanced.

e) The appraisal system can be used to protect the rights of individuals and provide due process for personnel decisions (fair and open criteria for appraisal, equity of treatment, equal rights or avoidance of discrimination, rights to negotiate and appeal). This use of appraisal is based on assumptions of validity in the appraisal system and open communication of performance standards and criteria. Open communications ensure that everyone is fully aware of the relevant standards of performance and will not be subject to arbitrary or biased standards. Validity ensures that people are appraised on their actual level of performance rather than on imagined or prejudged levels of performance. Such a system would not allow inadequate performers to hide, but more importantly would give recognition to competent performers who might otherwise not be recognized. These features of an appraisal system would go a long way towards ensuring equal rights and avoidance of discrimination in an organization.

f) The appraisal system can be used for professional development of groups and individuals in the organization. For individuals, this might come from feedback on their performance and discussions on how to improve (or what to continue doing) in the future. These discussions might even generate a set of action plans or personnel goals on which groups are to be appraised in the next evaluation. Training development needs for groups may not necessarily be identified through performance appraisal, but group progress toward some development goals could be assessed by aggregating some of the appraisal information. Development of individuals can also occur as a result of some of the other uses of appraisal such as learning the organization's preferred style of professional service, or having the opportunity to participate in goal setting.

Of course, not every appraisal system will use all these potential applications. This will be dependent on the needs or priorities of a particular organization. For example, if a school district decides to adopt a particular approach to teaching (i.e., the "Paideia" experiment in the City of Chicago school system),[176] then the appraisal system could concentrate on the areas of operating methods and organization goals needed to implement the new approach. If a school district operates as a "loose coupled" system,[177] relying on the delegation of general teaching responsibilities to teaching professionals, then the appraisal system will not be very complex and will concentrate on the area of personnel administration. If the school system is undergoing rapid growth and has to absorb a lot of inexperienced new teachers, then the appraisal system will be concerned with the areas of supervision and operating methods.

Within the same organization, different purposes or priorities for the appraisal system may exist for different levels in the hierarchy or different types of work in the organization. For example, certain professionals, such as consultants and specialists may have to work in a quite independent fashion with methods only known to themselves. In this case, an output-oriented or results-based appraisal system might be the most appropriate. Similarly for principals and supervisors, a goals-and plans-based system might be most appropriate with scant attention paid to the methods used by a principal. Presumably, screening and training prior to appointment would ensure confidence in the person's managerial abilities. For inexperienced classroom teachers, the appraisal system might focus on examining methods and skills while for experienced teachers, the evaluations might focus on teaching goals and plans. In all cases, an appraisal system will have to satisfy the requirements of due process. Most likely the concerns about due process will be safeguarded by unions or professional associations. But some may have very little experience with formal appraisal systems. In these cases, all parties would have to work together to ensure that the development of any sort of appraisal system satisfies requirements for due process.

---

[176] R. Stengel, "Quality, Not Just Quantity: The Paideia Proposal Aims to Reform America's Schools", Time, September 6, 1982, p. 87.

[177] K. W. Weick, "Administering Education in Loosely Coupled Schools", Phi Delta Kappan, June, 1982, pp. 673-676.

Once the priorities and purposes of appraisal for different parts of the organization have been determined, then attention can be drawn to the methods of appraisal. It is clear from the review of both sets of literature that methods of appraisal is a complicated and frustrating topic because all methods involve various compromises of cost, convenience, and complexity. The requirements for good appraisal methods are generally well understood (see section 6.1. on characteristics of effective standards), but the requirements are so demanding as to be practically impossible to satisfy by any one method of appraisal. Consequently, we are led to reiterate the earlier conclusion in section 10 of this report, which is that an organization must use a multifaceted approach.

While the multifaceted approach is also complicated, it is possible to reduce some of the complexity by using different methods of appraisal for different types of teachers or different levels in the hierarchy of a school. By way of illustration, exhibit C-6 presents a hypothetical example of how appraisal methods might vary in one school for different types of teachers. Each of the various methods in a multifaceted system must satisfy the criteria of validity, reliability and acceptability. These criteria can actually be enhanced by using different appraisal methods for different types of work. Moreover, a multifaceted approach avoids the problem of burdening everyone in the hierarchy appraisal methods which are inappropriate for some types of work.

For example, it may be futile to require a novice teacher to go through an elaborate goal planning activity for appraisal when the primary concern is whether the teacher possesses minimum acceptable standards of knowledge and skills for teaching in a subject area. Similarly, for an experienced teacher with a long history of excellence in teaching, it may be futile to force the teacher and the principal to go through a charade of detailed goals planning when the primary concern is to show recognition and appreciation for good work.

Exhibit C-6

Multifaceted Appraisal System: A Hypothetical Example

Type of Teacher	Focus of Appraisal	Evaluation Criteria	Possible Methods of Appraisal
1. Entry level or probation	decisions on retention or training	skills, traits knowledge, potential	credentials; rating and behaviour scales; incidents; simulations; possible ranking
2. Experienced permanent staff	maintain standards, professional development	continued teaching effectiveness, non-classroom work, development needs	observations; peers; secondary indicators; activities; negotiated plans
3. Senior teaching staff	maintain standards, wider responsibility	continued teaching effectiveness, administrative or leadership behaviour, results of plans	observations; negotiated work planning and review; results-focused appraisals
4. Special performance review	employee assistance, upgrading performance standards, dismissal	specific to problems, symptoms or indicators, specific timing and activity plans	negotiated plans, activities and targets; specific observations; trial periods; due process mechanisms

On the other hand, for a senior teacher who wishes to qualify for higher levels of responsibility, a detailed career development plan would be quite appropriate both for professional development and for performance appraisal. Thus a contingency approach seems to be the most suitable in the choice of appraisal methods. This in turn requires that senior administrators in a school system have a systematic knowledge of strengths, weaknesses and applicability of different appraisal methods.

Finally, another important consideration is the cost and complexity of administering an appraisal system. As was stated previously in section 10, many business organizations in the past have opted for simplistic methods of performance appraisal in order to go through the motions of appraisal as quickly and cheaply as possible. While all the faults of this approach are fully recognized, it is also recognized that virtually any system which tries to satisfy all the demands on a performance appraisal system will be very complex and cumbersome to develop and administer. This will be true for a multifaceted approach to appraisal as much as for any other modern approach to appraisal. However, a type of multifaceted approach can be used which can reduce some of the complexity and administrative burden of an evaluation system.

To reduce some of the complexity, it is necessary for an organization to define or establish certain minimum acceptable standards of performance (i.e., a respectable average level of performance) and develop some simple indicators of whether a person's performance is in this range. If the primary concern of the organization is to ensure that the great majority of professionals are performing at or better than the acceptable standard of performance, then no other form of appraisal is needed. This would be a maintenance approach to management, much like setting a ship on a course and then only making minor corrections to the steering to remain on course. Other, more complex forms of appraisal would be used for those persons whose performance is not yet up to, or has fallen below standard, or for those persons who wish to go through professional development to a different area. Thus an array of appraisal methods would be available for use in a school system; a "Pareto's optimum" approach could be used for the 20% of staff requiring special appraisals and simpler methods could be used for the 80% of staff requiring routine appraisals.

So far, the conclusions presented have been similar for the general management and the teacher evaluation literature on performance appraisal. We now turn to some of the major differences between the two sets of literature.

a) Career progression vs. organization tenure. While employees in a business organization must perform up to some level of acceptable standards to retain their jobs, there is a strong emphasis in such organizations on being able to qualify for higher level positions. The culture of expectations in such organizations is such that one must expect to have vertical career progression in the organization hierarchy. Thus the appraisal system must have a dual focus; giving feedback on current or past performance and also making assessments of potential for different kinds of work in the future.

In a public school organization, there is far less expectation of vertical career progression and instead the emphasis is on maintenance of organizational tenure or a culture of continuing to perform professional service work for the duration of a career. This is typical of professional service bureaucracies and the appraisal system consequently focuses essentially on the performance of professional service work, without any particular reference to potential for different kinds of work. Some professional bureaucracies do provide dual career ladders to a certain extent, to allow for career progression within professional service work without a person having to join the administrative hierarchy for career progression. However, even in these cases, the emphasis in performance appraisal continues to be on how well the person performs in their professional service tasks.

b) Output vs. process appraisals. Business organizations tend to have a greater use of output criteria in appraisal systems for a number of reasons. One is that many businesses have the opportunity to count or measure outputs in a variety of ways due to the nature of their production processes. Secondly, most businesses use some form of management by objectives to manage business operations and it is a natural extension of this to use an MBO form of performance appraisal. Thirdly, there is less concern with how results are achieved (save for legal and ethical constraints) than for what is achieved in business organizations. This may be due to lack of confidence in any one style or process for best obtaining results or due to a belief that many alternative processes are possible for achieving the same ends. In any case, the preference is to focus appraisals on results more than process.

In education organizations, the dominant focus of appraisal systems is on process rather than on outputs. This too is consistent with professional bureaucracy theory but also may be due to the difficulty (or reluctance) of measuring or counting outputs. Perhaps this is a function of public schools being indirectly financed by public tax resources because privately financed schools seem to be much more willing to use output-based performance appraisals. The widespread use of process-based appraisals is also interesting in the light of Levin's (1979) comment that so little satisfactory knowledge exists about which teaching methods do the best job of education.[178]

c) Performance-based vs. seniority-based compensation. Consistent with use of output-focused appraisals is the use of appraisals for determining pay differentials according to performance differentials. Most business compensation systems provide some universal cost-of-living pay adjustment, but rate of progress through a job salary range or progress to a higher salary range is determined by assessment of performance. Consequently, the appraisal system in some way must be concerned with ranking or comparing one individual's work to another. Similarly, business organizations are far more concerned with justifying compensation dollars and therefore are far more prone to dismiss unsatisfactory performers. This causes the appraisal system to have to rank-order performance evaluations in some manner.

In school organizations, the emphasis seems to be on acquiring entitlement to compensation adjustments by virtue of seniority and maintenance of adequate performance. Pay raises tend to be uniform with no performance differentials and progression through a salary range is by accumulation of years of seniority. Consequently there is no need for the performance appraisal to rank-order performance or otherwise distinguish between levels of performance except to identify serious cases of below-standard performance.

d) Non-union vs. unionized personnel. Few people at the professional service level in business organizations belong to bargaining units although they may belong to professional service associations. Consequently, their employment contracts or conditions of employment are based on uniform company personnel policies and standards plus any additional employment conditions which they may be able to negotiate. Performance appraisal will follow a similar pattern: a universal floor of common standards and practices for all plus individuated aspects of appraisal based on the person's unique circumstances and role in the organization. In practice, there is not a lot of variation in the overall performance appraisal system used for different individuals, but the principle of individuation exists for non-union personnel.

---

[178] Levin, "Teacher Evaluation," pp.4-6.

In teaching organizations, virtually all certificated personnel are members of bargaining units. A basic union philosophy is equal rights and treatment for all members. Consequently, the nature of the appraisal system tends to be universalistic as well. This is consistent with the approach to compensation discussed in section (c) above. Individuation in teacher evaluation does exist, of course, depending upon the strengths and weaknesses of individuals and their relationships with principals, but the influence of the union is to try to maintain universal equivalence as much as possible.

e) Personal vs. non-personal accountability for results. In business organizations, an operative assumption in performance appraisal is that individuals are essentially responsible for the achievement of assignments or agreed-upon results. They are given a certain amount of resources and a certain amount of time and are expected to be the catalyst which produces the outcome. Many results-focused systems have provision for extenuating circumstances, unforeseen events, etc., but ultimately the individual must accept personal accountability for the results produced in their domain of responsibility.

In contrast, it has been strongly argued in the literature on teacher evaluation that individuals cannot be held accountable for results in the public education system because of too many uncontrollable factors or extenuating circumstances. Instead the teacher is seen as just one of a series of actors in an educational environment which ultimately produces some form of education. Without discussing the merits of this view of accountability, it is quite clear that the viewpoint on personal accountability for results is quite different than that found in business organizations.

f) Internal vs. public accountability. In business organizations, the conventional wisdom is that boards of directors or senior executives accept the public accountability for the performance of their organization in society. Rarely, if ever does the public try to be involved in the internal operations of the organization except through communication with senior management. This means, in general, that accountability for individual performance within the organization is an internal matter between supervisors and subordinates. The work of individual professional service workers is less subject to public scrutiny and if the public does have some concern about the quality of this work, the matter is taken up with management rather than with the professional (an exception would be a direct fee-for-service relationship such as with a doctor or lawyer).

In public schools, there is far greater public scrutiny of the work of individual teachers and the public is equally as likely to confront the individual professional as the management of the organization. This is true even though there is no direct fee-for-service relationship between the public and the professional. The impact of this difference in concepts of accountability is that teacher appraisals have to be more sensitive to public accountability than would be true in most business organizations. This raises an aspect of complexity in teacher evaluations which would not be generally found in business organizations.

g) Whole vs. narrow performance evaluations. A "whole" job orientation in performance evaluation refers to the concept of evaluation on the complexity or multidimensional aspects of a job. If a job is relatively rich in content (i.e., extensive decision-making, autonomy, responsibility, multiskilled, etc.), then the appraisal system ought to be sufficiently multidimensional as to be able to reflect or capture the richness of the job. If the job is relatively low in complexity or essentially unidimensional, then the appraisal system can be correspondingly unidimensional or focus on a narrow, but important aspect of the job.

Despite the tendency in many business organizations to try to find a relatively simple and painless appraisal system, there is little tendency to try to focus an appraisal system on just one aspect of work. Business organizations seem to recognize that successful performance is a multidimensional phenomenon and there is an ongoing search for appraisal systems which can adequately reflect that complexity. In educational organizations, there is also a recognition of the complexity of a teacher's job, but the common practice is to focus appraisals more narrowly on classroom performance. Thus appraisals in schools tend to be more narrow in orientation while appraisals in businesses tend to be more whole-job oriented.

This situation does not pose a problem if the narrowly focused appraisal is concerned with a dimension of performance that is crucial to the organization or is a good indirect indicator or surrogate for the balance of the professional's work. For example, one might say that there is more to being a teacher than teaching in the classroom. However, if classroom performance was a valid representation of a person's overall performance in the school, then it would be sufficient to rely on the relatively narrow focus of evaluating classroom behaviours to make inferences about other dimensions of a teacher's performance. For instance, that good classroom behaviour indicates good relations with parents, good planning of lessons, or good grading of assignments.

If these assumptions are not valid and the narrow aspect of a job which is the focus of evaluation is not a good indicator of other important dimensions of performance, then the appraisal system may be of dubious validity. For example, if classroom behaviour is independent of other important teacher behaviours, then a narrowly focused appraisal system does not really reflect the wholeness of a person's work. This situation would be exacerbated if a narrowly focused appraisal system unfortunately evaluated performance on some relatively unimportant dimension of work such as neatness or punctuality that was easy to measure, but of inconsequential validity. The question then for narrowly focused teacher evaluations is twofold: which are the key dimensions of performance on which evaluations should focus, and are these dimensions adequate indicators of other important aspects of a teacher's work? The alternative is to develop multidimensional appraisal systems similar to business organizations which try to capture the complexity of the work of a professional.

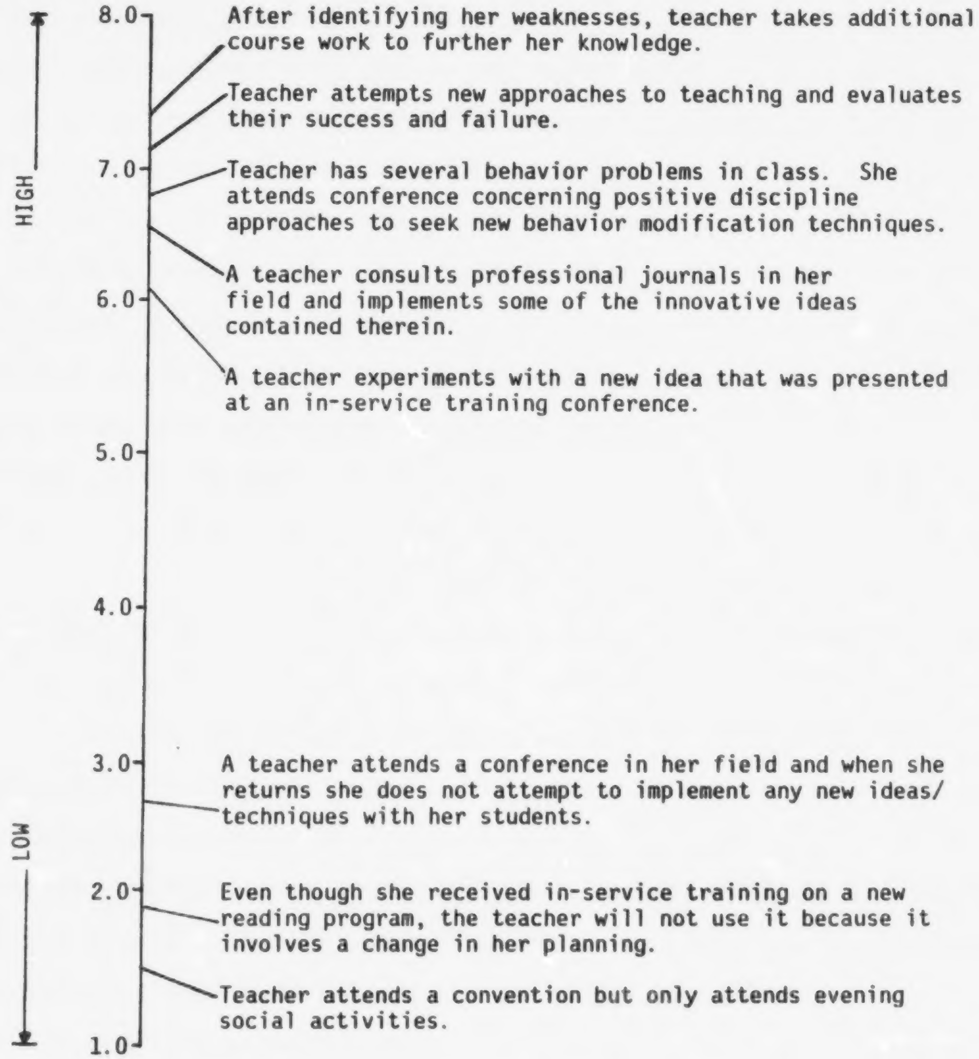
In summary, all of the above points suggest a series of cultural or philosophical differences which exist between performance appraisal in business and teacher evaluation. These are differences which have evolved over time and reflect the different nature of operations in the two types of organizations. However, the differences are not immutable and some possibilities for convergence of practices can exist. These convergences should not be sought merely for the sake of convergence. Appraisal practices from the business world should only be borrowed when they make sense to the needs and priorities of a particular school organization or for a particular segment of the certificated staff. Similarly, historical precedents in teacher evaluation should only be retained or perpetuated if they continue to make sense or meet the needs of a particular school system. Ultimately such judgements will have to be made by persons in each particular school system.

Appendix

Example of BARS For 8 Dimensions of Teacher Evaluation

1. PROFESSIONAL GROWTH

Seeks to improve instruction
through the exploration of suggestions,
new ideas, materials, approaches, etc.



After identifying her weaknesses, teacher takes additional course work to further her knowledge.

Teacher attempts new approaches to teaching and evaluates their success and failure.

Teacher has several behavior problems in class. She attends conference concerning positive discipline approaches to seek new behavior modification techniques.

A teacher consults professional journals in her field and implements some of the innovative ideas contained therein.

A teacher experiments with a new idea that was presented at an in-service training conference.

A teacher attends a conference in her field and when she returns she does not attempt to implement any new ideas/ techniques with her students.

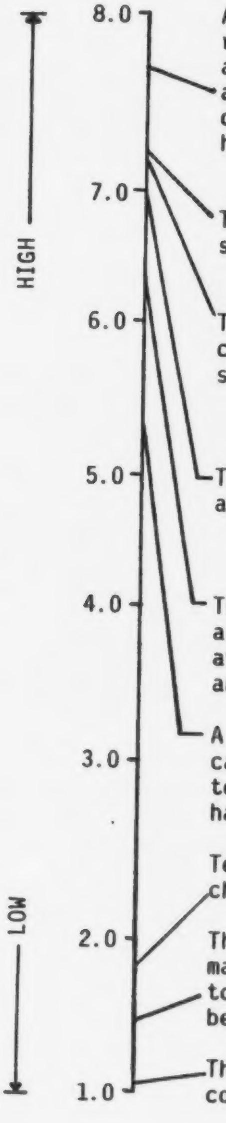
Even though she received in-service training on a new reading program, the teacher will not use it because it involves a change in her planning.

Teacher attends a convention but only attends evening social activities.

Reproduced from: Robert J. Beebe, "Use of the Behaviorally Anchored Rating Scale in Evaluating Teacher Performance" (Paper presented at the Annual Conference of the Evaluation Network, Tennessee, 1980).

2. COMMUNICATION SKILLS

The ability to exchange information with
administrators, professional persons,
parents, and other adults



A classroom teacher asks a speech therapist for information regarding the nature of a particular child's speech problem and how she can help. The therapist talks with the teacher about the problem and prepares a list of activities that the classroom teacher can implement to help the child improve his speech.

Teacher keeps in close contact with counsellors, or ancillary staff to keep all programs working toward the same goals.

The teacher holds regular conferences with parents, at their convenience. If parents are unable to come to the school, she makes a home visit.

Teacher consults with other professional staff in determining appropriate expectations for each child in the area of behavior.

The teacher confines her conversation with parents to her area of expertise, the classroom and instruction, and thus avoids treading into areas in which she lacks credentials and certification.

A child does not have his homework, so the teacher calls the mother and suggests that the child simply forgot to bring it to school, rather than accusing the child of not having done it at all.

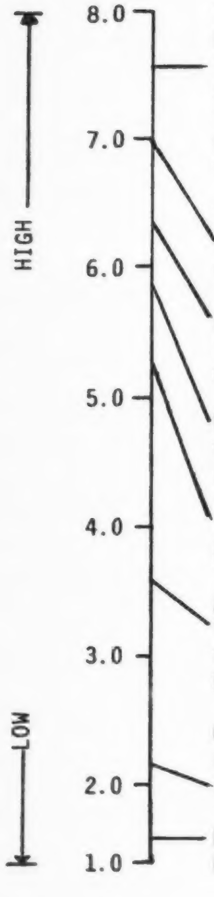
Teacher does not inform supervisor of major changes in a child's educational plan.

The teacher told the parents all year long that Johnny was making good progress. In the June conference, the teacher told Johnny's parents that she wanted to have him revaluated because his progress had been unsatisfactory.

The teacher ignores a parent's request for a conference.

3. RAPPORT WITH STUDENTS

The ability to develop an accepting relationship with children, which facilitates their success and growth in a learning environment.



The teacher observes that a child who has difficulty expressing his feelings in front of others is extremely upset about something. She also knows that this child is embarrassed to let the other students know he is upset. The teacher summons his help in running off and assembling dittoes in the office. This provides an opportunity for the child to talk to her.

The teacher extends to children the same common courtesies that she expects her children to give to her, e.g., please, thank you, excuse me, etc.

A student criticizes the teacher's style of discipline. The teacher listens to his opinion and tells the student why he has chosen this method of discipline.

When working with distraught adolescents who are having a rough time communicating with and understanding parents, a teacher relates her past experiences and problems in the same area.

The teacher avoids excessive familiarity with students so that he/she is perceived as the teacher and not a peer.

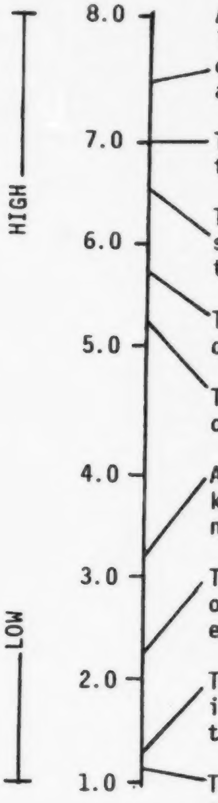
A teacher unknowingly mispronounced the child's last name. The child soiled himself daily for six weeks until the teacher finally said his name correctly, at which point the soiling of his clothing ceased completely.

Upon completion of instructional session, teacher criticizes student for making numerous errors, rather than focusing on what he completed correctly.

During instructional session, teacher remarks, "You'll never be able to learn this," when the student makes an error.

4. APPROPRIATE PLANNING

The ability to organize instruction in order to achieve short and long-term objectives based on the needs and level of performance of each child.



A language therapist reviews all testing data on a child and outlines her goals for the child. As her program progresses, she constantly revaluates her goals and objectives to make sure they are still appropriate to the child's needs.

The teacher is able to write an education prescription appropriate to the needs of each child.

The teacher has her lesson plan prepared and all materials that she will be using in a given lesson assembled and at hand before the lesson begins.

Teacher determines a student's entry level in a reading program on the basis of diagnostic test information.

Teacher organizes books, pictures, art materials for discussion on national holidays which occur during the school year.

A teacher has identified the long-term goals for the child. She knows generally what she wants to do to achieve these so she does not develop any short-term objectives.

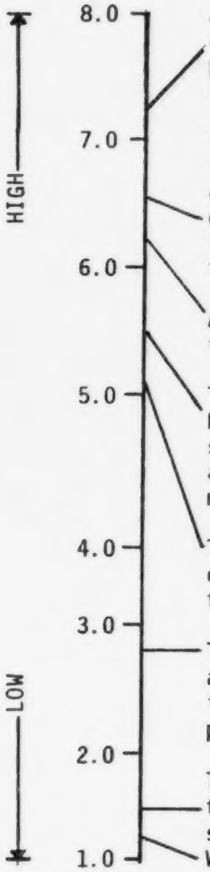
Teacher has not planned adequate amount of work to keep a student occupied and learning. As a result this student is bothering everyone else.

The teacher assigns the same dittoes as seat-work to all children in her class despite the fact that children's skills range from the first to the fourth-grade level.

Teacher establishes neither long-range goals nor short-term objectives for her students.

5. INSTRUCTIONAL SKILLS

The ability to implement the educational program
through the use of appropriate materials and techniques.



A teacher has commercial materials that are too advanced for her class. She modified them for use with individual students where possible. She also has numerous materials that she has made herself to fit individual needs.

In order to maintain the attention and interest of students, the teacher varies the pace of her instruction beginning with pencil-and-paper activities which demand much concentration on the part of the student and moving to a more relaxing activity such as an instructional game when she notices that the child's attention for the pencil-and-paper task is waning.

After a child is taught a particular skill, the teacher reinforces this skill.

The teacher is instructing a student in a second-grade level math book, but the student has not committed the basic addition and subtraction facts to memory. The teacher assigns student to an activity center where he practices these facts on a teaching machine.

The teacher uses the weekly magazine, News Pilot, to discuss current events and various social studies and science-related topics with her students.

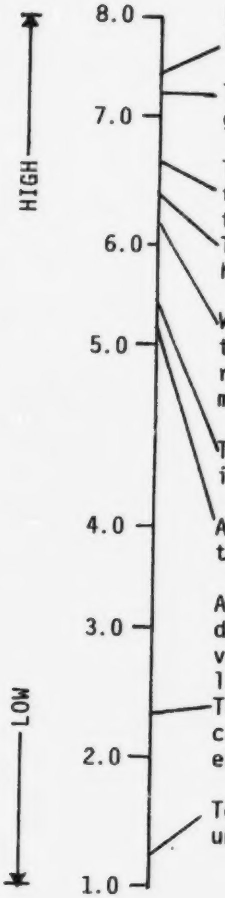
The teacher follows text in sequential order, making daily assignments to each pupil. Help is given each pupil as it is needed. No supplementary materials are used. No instruction per se is given.

The teacher uses audio-visual materials, e.g., films, tapes, to fill up time, with no thought being given to their relationship to the instructional program.

While teaching beginning reading, a teacher uses teacher-made dittoes on which the printing is very small and unclear.

6. RECORD KEEPING AND REPORTING

The ability to keep accurate and current records, using this information for comprehensive reporting as required.



Reports are always objectively written and substantiated by records; they are submitted within time requirements.

The teacher documents (in writing) the instructional goals and objectives she holds for each child.

The teacher gives a detailed account of when a child has a tantrum, including the time of the tantrum and the behavior that preceded the tantrum.

The special education teacher documents the communications she has with regular class teachers about special education students.

When asked to justify Johnny's placement in a third-grade math text, the teacher is able to present Key Math Diagnostic Test results and the textbook placement tests to substantiate placement at the third-grade level.

The teacher keeps a daily attendance record, collects absence information, and hands in attendance report every two weeks.

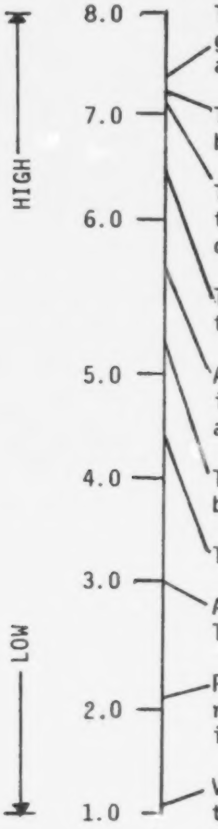
After administering various testing instruments to a student, the teacher places the results in the student's file.

A teacher claims she has "tried everything" to get Johnny to do his assignments but "nothing works." The teacher's supervisor asks to see a record of what has been tried, for how long, etc., so that she can offer direction to the teacher. The teacher has no written record of what has been done and can only give the supervisor vague recollections regarding each type of intervention that was tried.

Teacher writes reports which include emotionally laden unsubstantiated comments.

7. BEHAVIOR MANAGEMENT

The ability to establish and maintain a behavioral atmosphere that is conducive to learning.



The teacher established a behavior modification system involving group and individual goals for classroom behavior. Varying types and rates of reinforcement are established for individual students.

The teacher is alert to situations that can erupt into disruptive behavior. she takes steps to avert trouble before it occurs.

The teacher identifies student behaviors which interfere with the total learning process. She implements a behavior modification system utilizing reinforcers appropriate to each student.

The teacher uses a token economy with child who does not respond to social reinforcement.

A student has been evidencing disruptive behavior in the classroom in the form of calling out and interrupting. The teacher initiates a system for taking turns in speaking and asking questions.

Teacher rewards positive behaviors of children, ignores negative behavior.

The teacher takes away privileges to control classroom behavior.

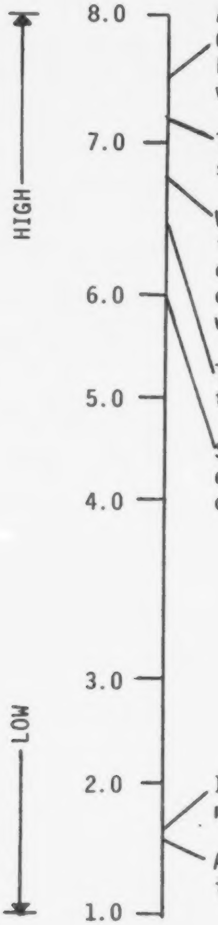
A teacher yelled at a student in a classroom because he misbehaved. The child ran out of school.

Pupils frequently swear and engage in name calling. The teacher reprimands them each time this occurs. No decrease in the inappropriate verbal behavior is noted.

Verbal and physically aggressive behavior frequently occurs in the classroom. The teacher provides no consistent consequences to these behaviors.

8. INTERSTAFF RELATIONSHIPS

The ability to build and develop positive relationships with I. U. and/or district staff.



A hearing therapist volunteers her services to a school on an Open School Night for Parents program by offering to speak, test hearing or whatever services the principal or other teachers may want of her.

The teacher volunteered to "serve" an extra duty for a regular staff member who became ill.

When the principal gave permission to have a ditto machine placed in the speech therapy room, the therapist said, "I would like to do my job effectively. It is impossible to conduct speech therapy classes effectively when the ditto machine is being used. Let's work out a schedule for use of the ditto machine."

The I. U. special class teacher adheres to the time schedules and the policies of the school in which she works.

Special class teacher maintains good public relations with regular class teachers by eating lunch with them and conversing with them on a personal as well as professional level.

I. U. Self-contained teacher in district building does not follow normal channels of communication within the building.

An itinerant teacher fails to inform the school office that she is ill and will be absent.

REFERENCES

Alberta. Alberta Education. Formal Evaluation of Certificated Education Staff in Alberta: Practices and Policies. Edmonton: Alberta Education, 1980.

Allen, P. M. et al. Teacher Self-Appraisal: A Way of Looking Over Your Own, Shoulder. Worthington, Ohio: Charles A. Jones, 1970.

Anderson, Scarvia B. Encyclopedia of Educational Evaluation. San Francisco: Jossey-Bass, 1975.

Appel, Joseph M. "Yearly Evaluation -- A Key in Assessing [the] Superintendent". Thrust For Educational Leadership, January 1980, p. 28.

Arikado, Marjorie and Musella, Donald. "Toward an Objective Evaluation of the School Principal", OCLEA, January 1975, pp.13-15.

Beebe, Robert J. "Use of the Behaviorally Anchored Rating Scale in Evaluating Teacher Performance". Paper presented at the Annual Conference of the Evaluation Network, Tennessee, 1980.

Beecher, Russell S. ."Staff Evaluation: The Essential Administrative Task". Phi Delta Kappan, 60 (1979), pp. 515-517.

Beegle, Charles W. and Brandt, Richard M., eds. Observational Methods in the Classroom. Washington, D.C.: Association for Supervision and Curriculum Development, 1973.

Berardi, Robert J. "Evaluation of Educational Personnel in the Commonwealth of Massachusetts". In The Improvement of Local Staff Evaluation and Supervision, edited by Robert J. Munnelly, pp.1-4. Massachusetts Association for Supervision and Curriculum Development, 1981. (ERIC Document Reproduction Service No. ED 208 525).

Borich, Gary D. and Fenton, Kathleen S., eds. The Appraisal of Teaching: Concepts and Process. Reading, Mass.: Addison-Wesley, 1977.

Borich, Gary D. and Madden, S. K. Evaluating Classroom Instruction: A Sourcebook of Instruments. Reading, Mass.: Addison-Wesley, 1977.

Bortner, Doyle M. "Coping with Demands For Instructional Accountability". NASSP Bulletin, May 1979, pp. 33-42.

Brandt, Richard M. "Towards a Taxonomy of Observational Information". In Observational Methods in the Classroom, edited by Charles W. Beegle and Richard M. Brandt, pp.23-34. Washington, D.C.: Association for Supervision and Curriculum Development, 1973.

Bryant, Gerald and Haack, Frank. "Appraisal: Peer-Centered and Administrator-Centered". Educational Leadership, May 1977, pp. 608-612.

Campbell, D. T. and Stanley, J. C. Experimental and Quasi-Experimental Designs For Research. Chicago, Illinois: Rand-McNally, 1963.

Carlson, Robert V. and Park, Richard. "Teacher Evaluation: Relevant Concepts and Related Procedures", June 1976. (ERIC Document Reproduction Service No. ED 129 739).

Castetter, W. B. The Personnel Function in Educational Administration. New York: MacMillan, 1976.

Coats, William D. "Accountability in Education -- The Kalamazoo Plan". Paper presented at the Annual Meeting of the America Association of School Administrators, Texas, 1975.

Cohen, Phyllis. "Teacher Evaluation in a School System: An Overview". Education Canada, 19 (1979), pp. 26-31.

Cross, Ray. "What Makes an Effective Principal?" Principal, 1981, pp. 19-22.

Deal, Terrence E.; Dornbusch, Sanford M.; and Crawford, Robert A. "Villians as Victims: Evaluating Principals". Phi Delta Kappan, 59 (1977), pp. 273-274.

Doremus, Richard R. "What Ever Happened to Kalamazoo's Merit Pay Plan?" Phi Delta Kappan, February 1980, pp. 409-410.

Educational Research Service. Evaluating Teacher Performance. Arlington, Virginia: ERS, Inc., 1978.

__________. Evaluating the Evaluator. Arlington, Virginia: ERS, Inc., 1973.

Ellett, Chad D.; Capie, William; and Johnson, Charles E. "Assessing Teacher Performance". Educational Leadership, December 1980, pp. 219-220.

Ertis, Bernie P. A. "How Much Are You Really Worth?" OSSTF Forum, December 1980, pp. 203-205.

"Evaluation: Confusion, Controversy Abound as Taxpayers Take Aim at Teachers". The Globe and Mail, February 23, 1981, p. 1.

Evertson, Carolyn M. and Holley, Freda M. "Classroom Observation". In Handbook of Teacher Evaluation, edited by Jason Millman. Beverly Hills, California: Sage Publications, 1981.

Farquhar, Robin H. "How the Teaching Profession Measures Teaching Effectiveness". Paper presented at the Conference of Teaching in Medical School, Saskatchewan, 1978.

Feldvebel, Alexander M. "Teacher Evaluation: Ingredients of a Credible Model". Clearing House May 1980, pp. 415-420.

"Flunking the Teacher". The Toronto Star, January 13, 1979, p. C4.

Flynn, John J. "Evaluation of School Administrators". Comment on Education, October 1978, pp. 5-8.

Gephart, William, ed. The Evaluation of Teaching. Bloomington, Indiana: Phi Delta Kappa Centre on Evaluation and Research, 1974.

Glasman, Naftaly S. "A Proposed Structure For Evaluating Personnel For Decision-Making". Studies in Educational Evaluation Spring 1977, pp. 47-56.

Glass, Gene. "Teacher Effectiveness". In Evaluating Educational Performance, edited by Herbert Walberg, pp.11-32. Berkeley, California: McCutchan, 1974.

Goodwin, Jean. "Constraints on Teacher Education". Dawson, Georgia: Terrel County Board of Education, 1977. (ERIC Document Reproduction Service No. ED 147 299).

Griffith, Francis. A Handbook For the Observation of Teaching and Learning. Midland, Michigan: Pendell Publishing Co., 1973.

Gudridge, Beatrice M. Teacher Competency: Problems and Solutions. Sacramento, California: American Association of School Administrators, Education News Service, 1980.

Haefele, Donald L. "How to Evaluate Thee, Teacher -- Let Me Count the Ways". Phi Delta Kappan, January 1980, pp. 349-352.

Hammons, A. L. "How To Evaluate Teachers on Performance". Paper presented at the annual meeting of the National School Boards Association, April 1975.

Hayman, John L. Evaluation in the Schools: A Human Process For Renewal. Monterey, California: Brooks, Cole Publishing Co., 1975.

Herman, Jerry J. "An Up-Front Approach to Administrative Evaluation". NASSP Bulletin, November 1978, pp. 34-42.

__________. Developing an Effective School Staff Evaluation Program. West Nyack, New York: Parker Publishing Co., 1973.

Holley, Freda M. and Hickman, Randall C. "Research on Teacher Evaluation: Needs and Realities". Paper presented at the annual meeting of the American Educational Research Association, Los Angeles, April 1981.

Horning, Alice S. "Teaching as Performance". Journal of General Education. Fall 1979, pp. 185-194.

Humphries, Kenneth W. "This Evaluation Plan Lets You Know What Your Administrators' Jobs are Worth". The American School Board Journal, May 1981, pp. 32-36.

Hunter, Madeline. "Teacher Competency: Problems, Theory, and Practice". Theory Into Practice, April 1976, pp. 162-170.

Johnson, Susan Moore. "Performance-Based Staff Layoffs in the Public Schools: Implementation and Outcomes". Harvard Educational Review, May 1980, pp. 214-233.

Joint AASA-NSBA Committee. Evaluating the Superintendent. Arlington, Virginia: American Association of School Administrators, 1980.

Jones, Meredith Howe; Jentz, Bary C.; and Wofford, Joan. "Anatomy of an Evaluation". Principal. September 1981, pp. 48-52.

King, B. and Jordan,. W. "Evaluation and Assessment of the Performance of Certificated Personnel". Fresno, California: Fresno County Department of Education, July 1972. (ERIC Document Reproduction Service No. ED 104 907).

Knoop, Robert. An Inventory of Competency Statements for Teachers, Principals and Superintendents. Toronto: Ministry of Education, Ontario, 1979.

Kowalski, Theodore J. "Why Your Board Needs Self-evaluation". The American School Board Journal. July 1981, pp. 21-23.

Lamb, Ronald W. and Thomas, M. Donald. "The Art and Science of Teacher Evaluation". Principal, September 1981, pp. 44-47.

Landrith, Harold E. "Systems Approach to Evaluation". Paper presented as the American Association of Colleges for Teacher Education Conference, Chicago, Illinois, 1977.

Latham, Gary P. and Yukl, Gary A. "A Review of Research in the Application of Goal Setting in Organizations". In Motivation and Work Behavior, edited by Richard M. Steers and Lymon W. Porter, pp. 488-509. New York: McGraw Hill Book Co., 1979.

Lehman, Irvin. "Evaluating Teaching". In The Evaluation of Teaching, edited by William Gephart. Bloomington, Indiana: Phi Delta Kappa Centre on Evaluation and Research, 1974.

Levin, Benjy. "Teacher Evaluation -- A Review of Research". Educational Leadership, October 1979, pp. 4-6.

Mallery, D. The Strength of a Good Faculty: Notes of Evaluation, Growth and Professional Partnership of Teachers. Boston, Mass.: National Association of Independent Schools, 1975.

Marks, J. R.; Stoops, E.; and King-Stoops, J. Handbook of Educational Supervision. Boston, Mass.: Allyn and Bacon, Inc., 1978.

Martin, Lawrence. "Teacher Competency Issue to be Tackled". ATA News, January 26, 1981, p. 5.

McCleary, Lloyd. "Evaluation of Principals". Theory Into Practice, February 1979, pp. 45-49.

McDonald, Frederick. "Principles and Procedures in Oberving Classroom Instruction". In Due Process in Teacher Evaluation, edited by Donovan Peterson and Annie Ward, pp. 110-113 Washington, D.C.: University Press of America, Inc., 1980.

McNally, Elaine F. and Birnie, Howard H. "Needs Assessment: Legislated or Voluntary?" The School Trustee, April 1978, pp. 8-19.

Melvin, J. Craig. "Performance Expectations: Teacher Supervision and Evaluation". Saskatchewan Education Administrator, December 1980, pp. 10-21.

Millard, Joseph E. "Teacher Evaluation". Ankeny, Iowa: Hartland Education Agency, 1976. (ERIC Document Reproduction Service No. ED 132 195).

Millman, Jason. Handbook of Teacher Evaluation. Beverly Hills, California: Sage Publications, 1981.

Mitsakos, Charles. "Generating a Valid and Reliable Rating Scale for the Assessment of Teaching Performance." In The Improvement of Local Staff Evaluation and Supervision, edited by Robert J. Munnelly, pp.10-14. Massachusetts Association for Supervision and Curriculum Development, 1981. (ERIC Document Reproduction Service No. ED 208 525).

Moberly, David L. and Stiles, Lindley J. "Getting A School Board To Address Its Primary Tasks". Phi Delta Kappan November 1978, pp. 236-37.

Munnelly, Robert J. "The Improvement of Local Staff Evaluation and Sypervision", Massachusetts Association For Supervision and Curriculum Development, May 1981. (ERIC Document Reproduction Service No. ED 208 525).

Nevo, David. "Toward a New Strategy for Evaluation Within The School Building". Urban Education, January 1979, pp. 435-452.

Newton, Robert R. "Teacher Evaluation: Focus on Outcomes". Peabody Journal of Education, October 1980, pp. 45-54.

O'Hanlon, James and Mortensen, Lynn. "Making Teacher Evaluation Work". Journal of Higher Education, November-December 1980, pp. 664-672.

Ornstein, Allan C. Accountability for Teachers and School Administrators. Belmont, California: Fearon Publishers, 1973.

Peterson, Donovan and Ward, Annie, eds. Due Process in Teacher Evaluation. Washington, D.C.: University Press of America, Inc., 1980.

Popham, James. "Performance Tests of Teaching Proficiency". American Educational Research Journal January 1971, pp. 105-17.

__________. "Pitfalls and Pratfalls of Teacher Evaluation". Educational Leadership, November 1974, pp. 141-146.

Potter, David; Nolan, Paul; and Lewardeski, Ann. The Relation of Student Achievement and Student Ratings of Teachers. Princeton, New Jersey: Educational Testing Service, 1973.

Redfern, George G. Evaluating Teachers and Administrators: A Performance Objectives Approach. Boulder, Colorado: Westview Press, 1980.

Renihan, F. I. "The Evaluation of Principals". Saskatchewan Education Administrator, December 1980, pp. 22-31.

Roelle, Robert J. and Wood, Philip A. "Six Ways to Improve Teacher Evaluation". The American School Board Journal, September 1980, pp. 36-41.

Roseavear, David. "Teacher Evaluation: Techniques and Tools". Comment on Education, February 1978, pp. 11-18.

Rosenshine, Barak and Furst, Norma. "The Use of Direct Observation to Study Teaching". In The Second Handbook of Research on Teaching, edited by Robert M. W. Travers, pp.122-183. Chicago, Illinois: Rand McNally, 1973.

Ross, Victor J. "Here's How Teachers Should be Evaluated". The American School Board Journal, August 1981, pp. 25-27.

Ryan, D. W. and Hickcox, E. S. Redefining Teacher Evaluation. Toronto, Ontario: Ontario Institute For Studies in Education, 1980.

Sandefur, J. T. An Illustrated Mode For The Evaluation of Teacher Education. Washington, D.C.: American Association of Colleges For Teacher Education, 1978.

Sapone, Carmelo V. "An Appraisal and Evaluation System for Teachers and Administrators". Educational Technology, May 1980, pp. 44-49.

Smith, Janice and Brown, T. J. "Relationships among Secondary Students' Evaluations of Their Courses and Teachers and their General Attitude Towards School". April 1976. (ERIC Document Reproduction Service No. ED 129 867).

Snyder, Karolyn J. "Clinical Supervision in the 1980s". Educational Leadership, April 1981, pp. 521-524.

Soar, Robert S. and Soar, Ruth M. Problems in Using Pupil Outcomes For Teachers Evaluation. Washington, D.C.: National Education Association, 1975.

Steers, Richard M. and Porter, Lyman W., eds. Motivation and Work Behaviour. New York: McGraw Hill Book Company, 1979.

Stengel, R. "Quality, Not Just Quanity: The Paideia Proposal Aims to Reform America's Schools". Time, September 6, 1982, p.87.

Stow, Shirley B. and Sweeney, Jim. "Developing a Teacher Performance Evaluation System". Educational Leadership, April 1981, pp. 539-451.

Sweeney, Jim and Stow, Shirley. "Performance Improvement -- A People Program". Education, Spring 1981, pp. 267-69.

Teacher and Administrator Evaluation. Bibliographies in Education. Ottawa, Ontario: Canadian Teachers' Federation, 1981.

Teacher Self-Evaluation Guide. Keewatin Region, Northwest Territories: Department of Education, 1980.

The Supervision of Teaching. Toronto, Ontario: The Canadian Education Association, 1979.

Tracey, William R. "Teacher Evaluation -- Another Perspective". Clearing Housing, January 1978, pp. 240-242.

Travers, Robert M. W. "Criteria of Good Teaching". In Handbook of Teacher Evaluation, edited by Jason Millman. Beverly Hills, California: Sage Publications, 1981.

________. Second Handbook of Research on Teaching. Chicago, Illinois: Rand McNally and Co., 1973.

Tymko, J. Lawrence. "Performance Appraisal: At Issue in the '80s". Alberta School Trustee, November-December 1980, pp. 10-11.

Walberg, Herbert, ed. Evaluating Educational Performance. Berkeley, California: McCutchan, 1974.

Walters, Donald L. "The Measurement of Administrative Competencies". Phi Delta Kappan, February 1980, pp. 423-25.

Weick, K. W. "Administering Education in Loosely Couple Schools". Phi Delt Kappan, June 1982, pp. 673-676.

Weigard, James, ed. Implementing Teacher Competencies. Englewood Cliffs, New Jersey: Prentice-Hall Inc., 1977.

Wolcott, Harry F. The Man in the Principal's Office: An Ethnography. New York: Holt Rinehart, 1973.

Wright, Robert. "Teacher Evaluation -- For the Improvement of Instruction", OCLEA, Fall 1979, pp. 3-6.

Zakrajsek, Barbra. "Evaluation Systems: A Critical Look". NASSPI, January 1979, pp. 100-111.

Zelenak, M. J. and Snyder, B. C. "Teacher Perceptions of the Teacher Evaluation Process". California Journal of Educational Research 25 (1974), pp. 205-212.

Volume

3

[illegible] ANALYSIS

[illegible]

[illegible]

[illegible]

A REVIEW AND ANALYSIS OF PERFORMANCE APPRAISAL PROCESSES
VOLUME III

Performance Appraisal for Professional Service Employees: Non-Technical Report

---

Professionalism in Schools Series

D. A. ONDRACK, Principal Investigator

C. OLIVER, Research Associate

This research project was funded under contract by the Ministry of Education, Ontario.

This study reflects the views of the authors and not necessarily those of the Ministry of Education.

The Honourable Sean Conway, Minister

© QUEEN'S PRINTER FOR ONTARIO, 1986

Order Information:

Publications Sales
The Ontario Institute
for Studies in Education
252 Bloor Street West
Toronto, Ontario M5S 1V6
Will invoice on orders over $30.00.
Other orders must be accompanied by a
cheque or money order payable to
O.I.S.E.

Publications Services
880 Bay Street, 5th Floor
Toronto, Ontario M7A 1N8

(416) 965-6015
(Toll Free) 1-800-268-7540
(Toll Free from area code 807)
Ask operator for Zenith 67200
Order must be accompanied by a cheque
or money order payable to the
Treasurer of Ontario

ONO 4189

Canadian Cataloguing in Publication Data

Ondrack, D. A.
A review and analysis of performance appraisal processes

Includes bibliographies.
Partial contents: v. 2. Performance appraisal for professional service employees, field study report -- v. 3. Performance appraisal for professional service employees, non-technical report.
ISBN 0-7729-1087-1 (3 v. set). -- 0-7729-1088-X (v. 1). -- 0-7729-1089-8 (v. 2). -- 0-7729-1090-1 (v. 3)

1. Teachers--Rating of. 2. Employees, Rating of. I. Oliver, Christine Tomaszuk. II. Ontario. Ministry of Education. III. Title. IV. Title: Performance appraisal for professional service employees.

LB2838.052 1986 371.1'44 C86-099627-1

TABLE OF CONTENTS

Page No.

Acknowledgements vi

Abstract vii

1. Introduction 1

2. Literature Review: A Summary 6
  - 2.1 Purposes of Appraisal 6
  - 2.2 Components of Appraisal 9
    - 2.2.1 Type of Criteria 9
    - 2.2.2 Method of Measurement 10
    - 2.2.3 Sources of Data Measurement 13

  - 2.3 Process of Appraisal 13
    - 2.3.1 Choice of Appraisal 13
    - 2.3.2 Appraisal Instruments 14
    - 2.3.3 Appraisal Interviews 17
    - 2.3.4 Appraisal Training 17
    - 2.3.5 Assessment Period and Timing 18
  - 2.4 Outcomes of Appraisal 19
  - 2.5 Link to Administrator Evaluation 20
  - 2.6 Comparison and Conclusions 22

3. Field Site Results 31
  - 3.1 Introduction 31
    - 3.1.1 The Business Group 31
    - 3.1.2 The Colleges Group 32
    - 3.1.3 The U.S. Schools Group 32
  - 3.2 Procedure 33
  - 3.3 Results 34
    - 3.3.1 Initial Design and Implementation of the System 34
    - 3.3.2 Current Objectives and Purposes of Appraisal 36
    - 3.3.3 Appraisal Design and Techniques 36
    - 3.3.4 Criteria or Standards for Appraisal 40
    - 3.3.5 Sources of Data on Performance 42
    - 3.3.6 Training and Preparation for Appraisal 45
    - 3.3.7 Special Cases and Organization Justice 48
    - 3.3.8 Evaluation and Maintenance of the Appraisal System 52
    - 3.3.9 Trends and Outlook for the Apraisal Systems 53

Page No.

4. Conclusions 53
  4.1 Objectives and Purposes of an Appraisal System 55
    4.1.1 Situational Factors 55
    4.1.2 Sense of Mission and Normative Models of Professional Service 56
    4.1.3 Career Stages of Professionals 57
    4.1.4 Absence of Mission or Normative Model 58
  4.2 Design, Implementation, and Support of Appraisal Systems 58
  4.3 Methods and Approaches to Appraisal 60
  4.4 Dimensions and Criteria for Appraisal 60
  4.5 Sources of Performance Data for Appraisals 61
  4.6 Standards for Evaluation of Performance 62
  4.7 Special Review and Renewed Probation 64
  4.8 Program Evaluation and Maintenance 64
  4.9 Summary Conclusions 65

Appendices 67

References 70

TABLES AND FIGURES

Page No.

Part 1

Exhibit 1: General Summary of Performance Appraisal Literature 3

Part 2

Exhibit 2: Choice of Appraiser: Advantages and Disadvantages 15
Exhibit 3: Multifaceted Appraisal System: A Hypothetical Model 25

Part 3

Table 1: Initial Design and Implementation of the Appraisal System 35
Table 2: Current Objectives and Purposes of Appraisal 37
Table 3: Specific Purposes of Appraisals 38
Table 4: Appraisal Design and Techniques 39
Table 5: Standards and Criteria of Performance 41
Table 6: Sources of Performance Appraisal Information 43
Table 7: Training and Preparation for Appraisals 46

Figure 1: Example of a Performance Monitoring System with Special Performance Review 50

ACKNOWLEDGEMENTS

The authors of this study wish to acknowledge the significant contributions of all field site personnel whom we interviewed for our study. Without their assistance and co-operation this project would not have been possible. Special recognition should be given to Dr. Glen Tecker whom we interviewed at Educational Testing Services, Princeton, New Jersey. His insights into the process and methods of appraisal were both inspiring and highly informative.

The authors would also like to acknowledge the guidance and assistance of the Advisory Committee of the Ministry of Education, under the chairmanship of Dr. Peter Nightingale, over the two years of work involved in this project. The comments and assistance received from the members of the Advisory Committee were a crucial asset to the progress and completion of the project.

ABSTRACT

This research project consisted of two major parts: a) a review of the literature on performance appraisal with special attention to the literature on certificated personnel in public education systems; b) field site studies of performance appraisal in action at ten professional service organizations in Ontario and seven public school systems in the USA. Each of the sites studied was selected on the basis of reputation for advanced practices in appraisals.

The literature review provided two major approaches to conducting performance appraisals: a) a focus on the "process" of professional work or the techniques, means, skills, and styles used to perform professional service; and b) a focus on the "outputs" of professional work, or various measures, outcomes, and judgements on the results of professional work. Secondly, the literature review provided two major uses of appraisals: a) "formative" uses where appraisal information is used for recognition, guidance, coaching, development, or problem correction in professional work; and b) "summative" uses, where appraisal information is used for administrative decisions on personnel such as compensation, promotion, training, transfers, or discipline. These approaches and uses of performance appraisal are not mutually exclusive and various combinations can be found in many systems.

Finally, another major theme in the literature review was the growing emphasis on protection of employee rights legislation and the vulnerability of employers to lawsuits arising from contentious personnel decisions based on performance appraisals. These trends have led to considerable effort to improve the reliability and validity of appraisal systems, but have also led to increased bureaucratization of procedures.

The field site studies mirrored many of the results from the literature review, with three important additions:

a) many appraisal systems are in constant evolution so that what is reported in the literature is sometimes not in practice when the site is visited;

b) different approaches and priorities in appraisal are used depending upon the local environmental situation of the organization;

c) bureaucratic cumbersomeness of many appraisal systems causes short-cuts to be used. In practice, this means that most teachers/professionals receive at least "satisfactory" ratings and the elaborate quasi-judicial safeguards built into many systems are used only rarely.

While no generally suitable appraisal system was found to emerge from either the literature review or field site study, two principles seemed to be operative in the most satisfactory sites:

a) under a set of general principles or guidelines established by a relevant central authority, a decentralized appraisal system is developed in operational detail by the local organization to reflect local concerns and priorities;

b) the decentralized system is developed jointly by supervisors and subordinates in the organization so that all involved parties accept the relevance and legitimacy of the appraisal system. Locally developed systems, which satisfy overall standards set by the larger organization, seem to be best suited to working through the inevitable learning curve of implementation and adaptation of an effective appraisal system.

1. INTRODUCTION

This report summarizes the findings of a contract research project authorized by the Ministry of Education of Ontario as one of a series of research studies on performance appraisal in the public education system. The purpose of this study was to conduct a review and analysis of published literature on performance appraisal in general and particularly on the use of appraisals in public education systems. Secondly, the study consisted of a series of field site investigations of performance appraisal systems in action to compare actual practice with results reported in the literature. Two types of field sites were chosen for the study: a sample of organizations in Ontario which employed a large proportion of professional service employees such as researchers, consultants, and educators, and a sample of public school boards in the United States which had a reputation for advanced practices in appraisal of teachers. The results from the field studies provided an opportunity to compare the information from the literature with appraisal practices in organizations which might be analogous to public school boards in Ontario. In this way, the vast range of information on performance appraisal could be narrowed or focused on information relevant to the needs of the public education system in Ontario.

In summary, the specific objectives of the study were:

i) to review the state-of-the-art theory and practice of performance appraisal approaches and methodologies;

ii) to identify those purposes of performance appraisal that might be appropriate to an educational system; and

iii) to describe those approaches which may be particularly relevant to the needs and practices of Ontario's educational system.

The results from this study are intended to provide a common base of information for use by the parties in Ontario most responsible for, and most affected by, performance appraisal of certificated educational staff. To insure that all parties shared the same perception of the nature of the study and the same access to results of the study, an Advisory Committee was established to oversee the definition and progress of the study. The Advisory Committee was composed of representatives from the Ministry, trustees, and professional associations involved in public education in Ontario, as well as representatives from other related contract research projects. Periodic progress reports by all the contractors were presented and discussed in Advisory Committee meetings to ensure co-ordination between projects and to obtain advice and commentary from the committee.

Research into the appraisal and evaluation of educational personnel reveals a broad-ranging and often complex set of issues that influence the acceptability and effectiveness of any system. Part two of this report presents a summary of the literature intended to highlight the individual and organization-level choices which must be exercised to implement or revise appraisal in any school system.

While the literature is replete with descriptions of the controversy, inconclusiveness, and complexity of teacher appraisal and evaluation, it is these features of performance appraisal which have helped to stimulate a body of research and commentary that offers the educator both a range of options in designing appraisal and greater sensitivity to the differing needs of participants to the process. The literature summary is organized into five categories of teacher performance appraisal: (1) *Purpose*; (2) *Components*; (3) *Process*; (4) *Outcomes*; and (5) *Link to Administrator Evaluation*. These categories are synthesized around the key questions that emerge in designing or revising an appraisal system. Exhibit 1 shows the five major topics and related sub-topics covered in part two of the report.

Exhibit 1

General Summary of Performance Appraisal Literature

DIMENSION	APPRAISAL ISSUE	KEY QUESTION(S)	DECISION OPTIONS
PURPOSE	Philosophy	. Why do we need to evaluate?	A. Accountability B. Improvement C. Supportiveness
	Objectives	. What do we specifically want to achieve with evaluation?	A. Formative Evaluation B. Summative Evaluation
COMPONENTS	Type of Criteria	. What makes a teacher effective?	A. Presage Criteria B. Process-Oriented Criteria C. Output-Oriented Criteria
		. Why is it so hard to define teacher effectiveness?	A. Complexity B. Context
	Method of Measurement	. How can we actually tell if the teacher is effective?	A. Personality Traits B. Performance/Behavior C. Skills/Competencies D. Teacher Tests E. Goal-Setting F. Clinical Supervision G. Goals-Free Evaluation
	Data for Measurement	. Where should the information for assessing effectiveness come from?	A. Classroom Observation B. Non-Classroom Activities
PROCESS	Choice of Appraiser	. Who should evaluate performance?	A. Supervisor B. Peers C. Self D. Students E. Expert Outsiders
		. How many people should evaluate?	A. Single Appraiser B. Multiple Appraisers C. Interactive Appraisal
	Appraisal Instruments	. What kind of documents should be used?	A. Checklists B. Narratives C. Rating Scales D. Goal-oriented Instruments
	Appraisal Interviews	. How should information be fed back to the teacher?	A. Process of Interviewing
	Appraisal Training	. Should training in the conduct of appraisal be provided?	A. Inservice B. Education Consultants
	Assessment Period and Timing	. How often should appraisal be conducted?	A. Observations B. Conferences C. Assessments
		. Should timing be fixed or based on need?	A. Scheduled Appraisals B. Appraisal in Response to Problems
		. Is evaluation a continuous or discontinuous process?	A. Day-to-Day Feedback B. Periodic Data Collection
OUTCOMES	Evaluation of Outcomes	. How do we evaluate the evaluation system itself?	A. Systematic B. Impressionistic
	Congruence of Outcomes with Stated Purposes	. Is appraisal doing what we wanted it to?	A. Acceptability B. Validity C. Satisfaction D. Accountability
LINK TO ADMINISTRATOR EVALUATION	Interactive Appraisal Design	. Should teacher appraisal be linked to administrator evaluation?	A. Separate Appraisal Systems B. System of Appraisal

Part three of this report presents results from field site studies of selected appraisal systems in actual practice. The field sites were grouped into three subgroups as follows:

i) a Business group composed of five profit-based or self-financing organizations providing professional services such as research, consulting, and teaching;

ii) a College group composed of five colleges of applied arts and technology (CAAT) in Ontario which are post secondary, tuition-based educational organizations; and

iii) a US Schools group composed of seven leading public school systems and one training and consulting organization for teacher evaluation.

All of the above organizations are analogous to Ontario public school systems in various ways and of course are different in various ways. These aspects of the field site sample will be discussed in greater detail in the section of the report which concentrates on the field site part of the research project. In general, the study of analogous organizations focused upon the following issues:

i) to identify the appraisal methodologies and the uses of appraisals in the various types of organizations;

ii) to identify major differences (if any) in approaches to appraisal which seem to be required by the different organizational types;

iii) to identify differences in approaches to appraisals in different functions of the organization and in different levels in the hierarchy;

iv) to identify whether process or output approaches to appraisals are typically used in some functions and organizations and not in others;

v) to identify various approaches to developing criteria and gathering data on the work of professionals; and

vi) to identify linkages between the appraisal and evaluation process and other parts of human resource management in the organization such as compensation systems, training and development, career planning, and supervision.

The final part of the report provides an integration of the results from the various investigations in the study as well as the authors' conclusions on the results. Since this study did not look directly at public education systems in Ontario (this was done in other related research projects), specific recommendations and conclusions for Ontario are not provided. Instead the authors have followed the requirement to provide a common information base to participants in the development of policy for the educational system in Ontario and it will be these persons who eventually develop specific conclusions and recommendations for use in Ontario. This report provides, to the best of our judgment, a set of generally applicable conclusions for an effective system of performance appraisal in education and it is hoped that these will provide a useful foundation for guidelines for future development.

2. LITERATURE REVIEW: A SUMMARY

2.1 PURPOSES OF APPRAISAL

The purposes of appraisal incorporate both the philosophy of the appraisal system and specific objectives. The philosophy of appraisal pertains to the question: why do we need to appraise and evaluate? Differing perspectives on the philosophy of teacher appraisal are implicit rather than explicit within the literature and uncover important beliefs, values, and assumptions surrounding the issue of performance appraisal. Among the most important of these are varying beliefs concerning:

(1) the significance of teacher effectiveness for the promotion of learning in relation to other variables (student effectiveness, principal effectiveness, school climate, policy, material support, for example);
(2) the potential for individual improvement within a supportive context;
(3) the degree of autonomy and self-determination necessary for effective teaching;
(4) the role of the principal in educational leadership and supervision;
(5) the potential for trust and interactive effort among different levels in the educational hierarchy; and
(6) the extent of co-ordination between instruction and administration in educational organizations that is necessary for student enhancement.

Although the literature does not explicitly address the notion of a philosophy of appraisal, it is revealed by recurring statements about the overall purposes of appraisal. These include:

(1) the promotion of personal and professional growth of employees;
(2) provision of a valid basis for making administrative decisions;
(3) improvement in instructional or school effectiveness;
(4) promotion of student progress and growth;
(5) improvement in the general quality of education; and
(6) accountability to stakeholders in the educational process (teachers, administrators, students, parents, the general public, and so on).

The importance of establishing a clear and acceptable philosophy of appraisal at the outset of system design is supported by two phenomena which pervade discussions in the literature:

(1) the agreement among most researchers and practitioners that some form of assessment is indispensable to the improvement of the quality of education; and
(2) the anxiety, discomfort, and, in some cases, resistance that currently attends the evaluation process for many teachers and administrators.

While there is fairly strong support in the literature for using evaluation, in general, to improve instruction, the specific objectives of appraisal are more susceptible to controversy.

The objectives of appraisal pertain to the question: what do we specifically want to achieve with appraisal and evaluation? Here, the literature tends to make strong distinctions between formative appraisals and summative evaluation. "Formative" appraisal refers to the use of appraisal for purposes of teacher development, feedback, and the identification of strengths and weaknesses. "Summative" evaluation, by contrast, refers to the evaluation of performance of individuals for purposes of making administrative or personnel decisions (e.g., employment, tenure, promotion, dismissal). Other sources of divergence identified in the literature, which address the difficulties of designing an acceptable appraisal system, are related to this distinction between formative and summative evaluation. These include:

(1) the paradoxical role of supervisor as judge versus helper;
(2) the issue of "accountability" versus "improvement";
(3) the "developmental" versus "evaluative" aspects of appraisal;
(4) the issue of teacher autonomy versus control and supervision; and
(5) the need to gather (and feed back) information on specific individuals versus the need to gather a broad sample of information for defining common standards among individuals.

Most of the educational literature speaks to the formative aspect of appraisals. The management and personnel literatures, by comparison, are more oriented to the use of appraisal for evaluation as well as motivational purposes.

The tendency for educational researchers and practitioners to de-emphasize summative evaluation has its roots in several factors:

(1) teacher behaviour is relatively autonomous and difficult to observe;
(2) the "outputs" or results of teaching are hard to assess;
(3) the teaching process requires considerable discretion and self-directed creativity in order to be effective;
(4) unlike summative evaluation, formative appraisal does not require the identification of common generalizable criteria for purposes of comparison among individuals;
(5) the organizational context of instruction (e.g., physical separation, high teacher-principal ratio) impedes informal performance feedback;
(6) personnel decisions in education organizations are often linked to seniority considerations;
(7) summative evaluation creates more anxiety among appraisers and appraisees than formative appraisals; and
(8) formative appraisal avoids the task of balancing the stated needs of various stakeholders in the education process, since formative appraisal is primarily teacher-focused.

The conceptual and practical division between formative appraisal and summative evaluation reflects a traditional division within the literature (and in practice) between the purpose of appraisal for improvement of performance versus judgment of performance. Recent literature, however, points to trends within the United States social and political system that underscore a growing emphasis on summative evaluation. For example:

(1) strong public response to the findings of the 1983 report of the National Commission on Excellence in Education;[1]
(2) growth in state and local school system adoption of teacher competency testing as a condition of certification;[2]
(3) increases in the pace of state-mandated teacher evaluation (in 1983, 26 states required teacher evaluation as compared to only six states in 1970); and
(4) public resurgence of the merit pay controversy.[3]

Although none of the current literature offers specific approaches for reconciling the formative-summative dichotomy, researchers and practitioners provide some general suggestions for combining the purposes of appraisal:

(1) teachers should be major participants in the design and process of appraisal;
(2) methods of appraisal should be part of in-service training for both administrators and teachers;
(3) day-to-day feedback and support needs to be an integral part of the overall appraisal system;
(4) one of the criteria in administrator assessment should be his or her effectiveness in appraising and evaluating teachers;
(5) the context of teaching should be understood so that teachers are only evaluated on the basis of behaviours under their control;
(6) the purposes and processes of appraisal and evaluation should be widely communicated (e.g., manuals, committees) and easily understood;
(7) some system of recourse or appeal must exist for complaints and grievances pertaining to the content and process of evaluation; and
(8) systems should be designed which are interactive, democratic, and mutually supportive rather than autocratic and control-oriented.

Overall, the literature assumes that resistance to and inconsistencies between formative appraisal and summative evaluation are minimal when the appraisal system is simple, valid, and equitable.

---

[1] See, for example: "Can the Schools Be Saved?" Newsweek, May 9, 1983.

[2] Hathaway, W.E. "Testing Teachers"; Educational Leadership Vol 33 (1980), pp. 210-215.

[3] See, for example, Toll, S. "This merit pay plan for management improves performance and sets an example for other employees (like teachers, for instance)". American School Board Journal Vol 170, no. 8 (August 1983), pp. 38-9; or Rist, M.C. "Our nationwide poll; most teachers endorse the merit pay concept". American School Board Journal Vol 170, no. 9 (September 1983), pp. 26-33.

2.2 COMPONENTS OF APPRAISAL

The components of appraisal can be summarized around three major topics:

(1) type of criteria;
(2) method of measurement; and
(3) data for measurement

2.2.1 Type of Criteria

Choice among types of criteria involves the question: *what makes a teacher effective?* Lack of conclusive evidence on teacher effectiveness is attributed to a host of factors in the research literature. In general, the difficulty of determining what makes a teacher effective pertains to two issues, *complexity* and *context*, both of which inhibit the identification of commonalities among teachers for the prediction of effectiveness. *Complexity* refers to the availability and exercise of diverse means (teacher behaviours, activities, and approaches) for the attainment of a single end (pupil progress). Most arguments in the literature surround two distinct aspects of complexity: the degree of complexity, diversity, or ambiguity *inherent* in the teaching process,[4] and the degree of discretion and flexibility there *should* be. While there is a general consensus within the literature that the teaching process is complex, the greater the number of appropriate teaching methods (that either *do* exist or *should be* permitted) for achieving the single end of student growth, the greater the difficulties in establishing common criteria that meaningfully distinguish between good and marginal teachers for purposes of appraisal. The second problem in identifying valid criteria of teacher effectiveness is *context*. The causal link between teacher effectiveness and student outcomes is moderated significantly by a large number of contextual factors; what is more, the presence and influence of these factors are likely to vary considerably among teachers. Context factors include:

(1) student characteristics (e.g., grade level, motivation, socio-economic status);
(2) instructional setting (e.g., materials and professional resources, structure of curriculum);
(3) school climate, policy and organization;[5] and
(4) the social context and conflicting expectations of different educational constituencies.

---

[4] See, for example, Darling-Hammond, L.; Wise, A.E.; and Pease, S.R. "Teacher Evaluation in the Organizational Context: A Review of the Literature". *Review of Educational Research* Vol 53, no. 3 (Fall 1983), pp. 285-328.

[5] See, for example, Lotto, L.S. "Revisiting the Role of Organizational Effectiveness in Education Evaluation". *Educational Evaluation and Policy Analysis* Vol. 5 no. 3 (Fall 1983), pp. 367-378; or Weick, K.E. "Administered Education in Loosely Coupled Schools". *Phi Delta Kappan* Vol. 63, no. 10 (1982), pp. 673-76.

All of these variables tend to limit the number of criteria of teacher effectiveness which are generalizable across educational settings.

There are three types of criteria for assessing teacher effectiveness:

(1) presage criteria: the characteristics with which the teacher comes to the job, including personal and professional characteristics;

(2) process-oriented criteria: criteria based on the process of teaching, including teaching methods, behaviours, and techniques, student-teacher relationships, and classroom strategies; and

(3) output-oriented criteria: criteria based on student learnings, pupil outputs, and changes in pupil behaviour.

A large number of advantages and disadvantages adhere to each of these types of criteria depending upon the objectives and content of the appraisal program. Nevertheless, some tentative conclusions from the literature appear warranted at this time:

(1) presage criteria are more applicable for use in the recruitment and selection of teachers than as standards to be utilized in ongoing assessment;
(2) presage criteria which correlate with teaching effectiveness are: scholarship in college, professional preparation, attitude toward teaching, age, and experience;
(3) presage criteria which appear uncorrelated with teaching effectiveness are: knowledge of subject matter, cultural background, intelligence, socio-economic status, sex, and marital status;[6]
(4) process-oriented criteria are more common and easier to use than output-oriented criteria; the use of output-oriented criteria tends to create resistance to evaluation; and
(5) the validity of output-oriented criteria suffers from the inability to control for all context variables that moderate the teacher behaviour-pupil learning relationship.

2.2.2 Method of Measurement

Whereas choice of criteria attempts to answer the question "what makes a teacher effective?" choice of method of measurement asks: how can we actually tell if the teacher is effective? Seven methods emerge from a review of the literature; these are summarized below.

(1) Traits (Personality Characteristics)

---

[6] See Farquhar, R.H. "How the Teaching Profession Measures Teaching Effectiveness". 1978 (ED 1985 683); or Feldvebel, A.M. "Teacher Evaluation: Ingredients of a Credible Model". Clearing House, May 1980, pp. 415-420.

Assessment of teachers on the basis of personality traits tends to produce measures which are ambiguous or non-specific (e.g., "the teacher is conscientious", or "enthusiastic"); this ambiguity makes the process of appraisal much more prone to subjectivity and appraiser bias. Also, because this type of measurement lacks behavioural specifics, teachers find it hard to determine how to improve their performance. Despite their continued use, assessment on the basis of traits is usually unreliable and inaccurate.

(2) Performance/Behaviour

This method of measurement examines the actions or behaviours which are expected to occur if performance is satisfactory. Performance based or behavioural data possess the greatest potential for validity, reliability, discriminability, and receptivity. The measurement of specific behaviours also allows for concrete and meaningful feedback to the teacher. The development of behaviour-based standards is more time-consuming than the trait approach, but existing instruments[7] can be borrowed and revised to accommodate individual district or school needs.

(3) Skills and Competencies

Unlike behaviour-based measurement of performance, assessment of competencies or skills only measures capacity. That is, skills-based assessment measures what the teacher is able to do rather than the extent to which this capacity is applied consistently or appropriately. This distinction is commonly overlooked in the education literature. Considerable progress has been made in developing a valid inventory of competency statements that describe effective teaching.[8]

(4) Teacher Tests

Although 15 U.S. states now require some form of tests for teachers as a condition of certification (e.g., The National Teacher Examinations), the use of teacher tests has not been applied to the appraisal of certificated personnel and is unlikely to be used as a basis of measurement in the future. Despite a Supreme Court ruling that NTE scores are related to teacher job performance, the validity of teacher testing is still a source of controversy in the literature.

(5) Goal-setting (MBO)

Unlike measurement against a predetermined set of criteria, goal-setting involves the assessment of performance against mutually derived goals and objectives. Research and commentary on the advantages and disadvantages of goal-based appraisal are numerous. For example, goal-setting tends to be time-consuming, provides limited feedback on specific teacher behaviours, and requires consistent commitment by all participants in order to be useful. On the other hand, goal-setting is an interactive process which encourages hierarchical communication, self-direction, self-evaluation, and co-operation between teacher and supervisor. There is no conclusive evidence at present on the acceptability and validity of goal-setting.[9] Several sources advocate its use in combination with predetermined criteria.

---

[7] For suggested sources, see Lewis, A.C. Evaluating Educational Personnel. American Association of School Administrators, 1982.

[8] See for example, Gutridge B.M. Teacher Competency: Problems and Solutions. American Association of School Administrators, 1980.

(6) Clinical Supervision

Clinical supervision covers a fairly broad literature that describes a democratic, interactive approach to the improvement of instruction. Clinical supervision consists of three major activities: planning conferences, classroom observation, and feedback conferences. The objectives of clinical supervision are professional development, objective feedback, skill development, and the diagnosis of instructional problems. Its use as a basis for making personnel decisions (summative evaluation) has been suggested but not explored in any detail in the literature. The links between clinical supervision and teacher performance, and between clinical supervision and student performance, have not yet been convincingly demonstrated.[10]

(7) Goal-Free Evaluation

A final method of measurement which departs significantly from the foregoing approaches is "goal-free evaluation".[11] This approach argues against the use of any specific, standard set of criteria. Instead, teachers "make their best case for merit or value" based on multiple lines of evidence that reflect the needs and understandings of all audiences who use the evaluation results. These "lines of evidence" may in fact constitute, among other things, several methods of measurement or types of criteria described earlier (e.g., performance tests, personal characteristics, credentials, competency-based evaluation). Deficient teachers are identified as those who are unable to provide credible or sufficient evidence to be considered to have an adequate case. The validity of this method of measurement has yet to be assessed. Although goal-free evaluation overcomes the problem of identifying commonalities among teachers for purposes of comparison and attempts to respond to the needs of all stakeholders in the process, this method is more vulnerable to evaluator bias and subjectivity in making summative assessments.

---

[9] For a thorough treatment of this approach, see Redfern, G.B. Evaluating Teachers and Administrators: A Performance Objectives Approach. Westview Press, Boulder, Colorado: 1980.

[10] Acheson, K.A., and M.D. Gall, Techniques in the Clinical Supervision of Teachers. New York: Longman, 1980.

[11] Peterson, K., and D. Kauchak, "Progress on Development of Lines of Evidence for the Evaluation of Public School Teaching". Paper presented at the American Educational Research Association, Montreal, Apr. 1, 1983 (ED 228 228).

2.2.3 Sources of Data For Measurement

The third component of appraisal concerns data for measurement. This aspect of appraisal is a response to the question: where should the information for assessing teacher effectiveness come from? There are two sources of appraisal data: classroom observation and non-classroom activities. The current literature focuses almost exclusively on the former data source and relates specifically to the "how-to" of observation.[12] Conceptualizations of alternative data sources can be re-cast in terms of who should do the appraising, since choice of appraiser (e.g., outsider, principal, teacher) is the primary determinant of what data are included in appraisal. This brings us to the first of several process factors in the appraisal of teacher effectiveness.

2.3 PROCESS OF APPRAISAL

The process of appraisal refers to the mechanics of how appraisal is conducted. There are five major aspects to the process of appraisal:

(1) choice of appraiser,
(2) appraisal instruments,
(3) appraisal interviews,
(4) appraisal training, and
(5) assessment period and timing.

2.3.1 Choice of Appraiser

Choice of appraiser pertains to two specific questions: who should appraise and evaluate performance? and how many people? A variety of sources in the literature address these issues either directly or indirectly.

(1) Who should appraise and evaluate performance?

Answers to this question are likely to evolve, in practice, from the implicit philosophy of the appraisal system itself. For example, an emphasis on development, growth, and autonomy requires self-evaluation as an integral component of the appraisal process. Choice of appraiser also emerges from the prescribed objectives of appraisal. Where the purpose of performance appraisal is summative, the supervisor is assumed to have responsibility for making the evaluations of the relative performance of teachers upon which personnel decisions will be based. Furthermore, the decision as to who should participate in appraisal is also influenced by the strength of interest and degree of perceived co-operation among different constituents in the education process (for example, the amount and acceptability of input from public officials, parents, and outside experts). Direct participants in the appraisal process may include: (1) supervisor; (2) self; (3) peers; (4) students/subordinates; or (5) expert outsiders. There is fairly extensive research and practical commentary on the efficacy of these alternative sources; a summary of advantages and disadvantages is outlined below in Exhibit 2.

---

[12] For a particularly thorough treatment of this topic, see McDonald, F. "Principles and Procedures in Observing Classroom Instruction". In Peterson, D., and A. Ward, eds., Due Process in Teacher Evaluation. Washington, D.C.: University Press of America, 1980, pp. 110-113.

All appraisal sources are prone to some degree of bias and no research to date has provided conclusive evidence on the relative empirical validity of these five sources. Nonetheless, investigations into methods for reducing rater error have been most vigorous in the area of supervisor appraisals, since this form of assessment continues to be the most common.

(b) How many people?

The question of how many people should appraise and evaluate is important for several reasons. First, the decision to use only one appraiser precludes the opportunity for making appraisal an interactive, mutually supportive process. In this way, the choice of appraiser becomes a manifestation of the philosophy which underlies the system. Second, research indicates that the use of multiple sources of evidence tends to increase the validity and reliability of the appraisal results and reduce judgment error or bias. For this reason, some schools experiment with a "team approach" to appraisal. The benefits of increased validity, however, have to be considered in light of two potential disadvantages: the use of multiple appraisers is more costly and time-consuming, and the process may be threatening to the appraisee. Finally, knowing to whom the appraisal data will be made available also affects the number of data sources chosen. For example, teachers are less resistant to formal input from students, outsiders, and peers if such information is fed back only to the teacher. The proposed audience for appraisal data is, in fact, a reflection of the purposes that the appraisal system is intended to serve. Formative appraisal does not require teachers to share data with others (although it is, naturally, encouraged). The validity of summative evaluation, by comparison, increases in proportion to the amount of data or number of sources available for making judgments about teacher effectiveness.

2.3.2 Appraisal Instruments

Appraisal instruments refer to the specific documents used in evaluating teacher effectiveness. There are only limited examples of actual appraisal instruments in the educational literature[13] but common instrument characteristics tend to vary among four overa designs. These are summarized below.

---

[13] See, for example, Gutridge, B.M. Teacher Competency: Problems and Solutions, op. cit.; Lewis, A.C. Evaluating Educational Personnel, op. cit.

Exhibit 2

CHOICE OF APPRAISER: ADVANTAGES AND DISADVANTAGES

Choice of Appraiser	Advantages	Disadvantages
SUPERVISOR	- supervisor hierarchically - accountable for the performance of subordinates - supervisor responsible for personnel decisions - provides a forum for supervisor-appraisee interaction and communication	- may promote discomfort for appraiser or appraisee - prone to bias and subjectivity
SELF	- suitable for developmental purposes - encourages self-growth and motivation - overcomes difficulty of ongoing observation by externals	- unsuitable for purposes of making personnel decisions - tendency toward bias and distortion - low predictive validity for student achievement
PEER	- fellow professionals provide relevant valuable input - peers more apt to take into account day-to-day job demands	- peer group pressures may create bias or distortion - may create disharmony or competition - timetable restrictions impede its practicability
STUDENT/ SUBORDINATE	- suitable for developmental purposes - provides valuable feedback to appraisee	- inconclusive evidence on predictive validity - age of student restricts its use
EXPERT OUTSIDER	- increases validity and objectivity of appraisal - provides valuable 'expert' feedback - useful for resolution of appraiser-appraisee conflict - saves supervisory time	- may be expensive - process may be threatening to appraisee - use of third party discourages meaningful interaction about performance improvement between supervisor and subordinate

(1) Checklists

Checklists are comprised of a list of statements (describing specific traits or behaviours) beside which the appraiser is asked to check which are applicable to a particular employee or observed by the appraiser. Checklists indicate the presence or absence of a particular attribute without reference to the degree to which that attribute prevails. During observation, a checklist is simple to use and directs attention to aspects of a lesson which an appraiser might otherwise miss. On the other hand, classroom observation may become a mechanical, routine procedure in which supervisors are apt to make judgements without patient reflection or careful analysis.

(2) Narratives

The narrative technique (or "essay evaluation") is a written report of the teacher's strengths, weaknesses, or future potential. The rich descriptive nature of this method makes it a worthy supplement to other types of techniques for purposes of feedback. Narratives are of limited use in isolation, however, as a basis for making judgemental decisions that involve meaningful comparison among teachers.

(3) Rating Scales

The most widely used evaluation technique involves judgemental rating of various characteristics of the teacher's performance along some graphic scale or continuum. The scale can range in specificity from a simple "unsatisfactory - needs improvement - satisfactory" to more specific descriptions of different levels of performance or exhibited behaviour (e.g., behaviourally anchored rating scales). Behavioral rating scales are more useful for appraisal-feedback purposes than simple judgemental rating scales which tend to be used more for evaluation purposes.

(4) Goal-Oriented Instruments

This type of instrument is used when the method of measurement for appraisal and evaluation is goal-setting. Although the format of the instrument depends on the individual designers or users, goal-oriented instruments usually provide space for one or more of the following: (1) identifying the goals to be achieved; (2) specifying how the goals will be accomplished; (3) describing or rating the extent to which goals have been achieved; (4) providing evidence of progress toward goals; and/or (5) identifying personal or organizational factors which impede or facilitate goal accomplishment. Goal-oriented instruments provide a rich amount of appraisal data which can also be used for evaluation purposes.

The choice of which technique to use is contingent upon its purpose (formative appraisal versus summative evaluation) and the appraisal designer's individual assessment of which benefits and drawbacks inherent in each technique are most relevant to the particular context.

An alternative approach is instrument design that combines more than one of the foregoing techniques. Because hybrid techniques overcome several of the weaknesses of any single approach used in isolation, instruments are sometimes designed, in practice, to include both goal statements and performance-based rating scales.

2.3.3 Appraisal Interviews

The choice of appraiser and type of instrument are concerned with where data should come from and how information should be collected. Appraisal effectiveness also depends on how the information is fed back to the teacher; a good appraisal interview (or conference) is both systematic and humane. To a large extent, the discomfort which often attends the appraisal interview is a function of the quality of (and participation in) decisions made on all other performance appraisal issues. For example, the appraisal interview is likely to be less stressful if:

(1) both the supervisor and teacher have a clear and compatible understanding of why the evaluation has been conducted (purpose of appraisal);
(2) the teacher knows what is expected of him or her (criteria and standards);
(3) the teacher expects that the process will be supportive (articulation of appraisal philosophy);
(4) the teacher knows in advance how the appraiser has assessed performance (data for measurement and appraisal instruments); and
(5) the teacher has full knowledge of where the information has come from (appraisal sources) and how it will be used (objectives of appraisal).

The educational literature suggests that while appraisal conferences are acknowledged as a key contributing factor to successful appraisal, little research has been conducted to guide practitioners through the process.[14] The management and personnel literatures, by comparison, have generated a broad range of advice and instruction on how to conduct an appraisal interview. Appendix 1 provides a synthesis of the bulk of the more recent literature on this topic. Overall, the keys to successful interviewing are careful preparation, open dialogue, and a problem-solving rather than fault-finding climate.

2.3.4 Appraisal Training

Inservice training in the process of performance appraisal (for both appraisers and appraisees) may be another major determinant of successful appraisal. To date, both research and practice have tended to neglect this issue. Yet there are several potential benefits of inservice training on the "how-to" of appraisals. For example, formal training programs may:

(1) increase both the appraiser's and teacher's confidence or faith in the appraisal process;
(2) clarify the purposes of appraisal;

---

[14] Sweeney, J. "The Post-Observation Conference: Key to Teacher Improvement". The High School Journal Vol 66, no. 2 (1983), pp. 135-140.

(3) provide a vehicle for disseminating information on appropriate appraisal methods;
(4) reveal unintentional rater biases;
(5) promote consistency and a common frame of reference for the conduct of appraisal; and
(6) encourage involvement and participation among hierarchical levels within the school organization.

While it is too early to assess the effectiveness of such endeavors, a small minority of schools and school districts in the United States are beginning to develop successful inservice training programs or hiring experts in the field of education to conduct workshops and seminars on the appraisal process.

2.3.5 Assessment Period and Timing

The fifth and final dimension of the appraisal process pertains to three distinct questions related to assessment period and timing. First, how often should assessments be conducted? There is currently no specific research on the optimal number of times per year that assessment should occur. The more often that data are collected, the more timely the information and, therefore, the more accurate. However, time, cost, and convenience considerations usually restrict the frequency of both observations and evaluations, in practice, to once a year or once every few years.

The second question which deserves attention is: should timing be fixed or based on need? Again, there is no explicit treatment of this question in the literature, but the answer, in practice, commonly reveals the stimulus which drives the evaluation process. That is, some appraisers adhere to a systematic schedule of observation, conferences, and summative assessments, whereas others only evaluate when a critical problem surfaces. In the latter case, supervision becomes "management by exception" and the timing of assessment is driven by the emergence of problems or negative information rather than the positive desire to assist all teachers. Under such circumstances, the supervisor may "waive" observations or formative appraisals for those teachers whom he or she perceives as performing adequately, which in turn denies good teachers the opportunity for positive feedback and sustains an image of the appraisal process as negative or punitive.

There is one final question which relates to assessment period and timing: is appraisal continuous or discontinuous? Whatever interval for assessment is chosen, appraisal can be defined as a continuous process, because information about employee performance is continuously monitored and feedback to the teacher should occur often on a more informal basis. Where the appraisal process is effective, the day-to-day support of the teacher should be a natural outgrowth of the formative appraisal system. Evaluations, however, may be a more periodic event as required by the summative evaluation needs of the system or by particular problems of performance.

2.4 OUTCOMES OF APPRAISAL

If an appraisal system is effective, the outcomes of appraisal are virtually identical to the purposes of appraisal (e.g., student growth, instructional effectiveness, professional improvement, valid personnel decisions, teacher feedback, and support). None of the current literature confronts the evaluation of system effectiveness specifically, that is, how do we evaluate the evaluation system itself? Impressions of effectiveness of appraisal systems tend to be inferred informally from one or more of several indicators; for example: satisfaction, school morale, public response, number of grievances, level of absenteeism, pupil test scores, informal feedback, amount of paperwork, or costliness of the system. One notable exception to the absence of systematic methods for evaluating the appraisal program itself is the Teacher Evaluation Needs Identification Survey.[15] Developed in Hartford, Connecticut, to determine the degree to which a school's evaluation system complies with state guidelines for effective teacher evaluation programs, TENIS comprises a staff survey that assesses the status of teacher evaluation practices in local schools. Results obtained from field tests substantiate the instrument's validity and reliability as well as its sensitivity to differences in teacher evaluation practices.

In general, the outcomes of appraisal concern the question: is appraisal doing what we wanted it to? In the educational literature, outcome issues are addressed within a more conceptual framework that relates to one or more of the following criteria of appraisal effectiveness:

(1) acceptability: the extent to which appraisal participants and outside constituents exhibit trust in, rather than resistance to, the process;

(2) validity: the extent to which the appraisal system actually measures what it purports to measure with minimal bias or error;

(3) satisfaction: the extent to which the system meets the needs of individuals and promotes a positive, meaningful working climate; and

(4) accountability: the extent to which the appraisal system clarifies individual and organizational roles and ensures prescribed responsibility for improvement in the quality of education.

Although these four outcome issues are usually treated separately in the literature, they are actually interrelated, in that: (a) acceptability is a function of the degree to which individuals are satisfied with the system and believe that it is valid; (b) the validity of the system is a function of the degree to which individuals are willing to accept the way that appraisal is conducted and the purposes it serves; and (c) satisfaction with the system is a function of the degree to which the process appears valid and equitable. Resistance to appraisal is less the result of an unwillingness on the part of individuals to be held accountable for their actions than the consequence of breakdowns in the conditions that promote accountability. Generally speaking, actions for which individuals are accountable need to be (a) clearly specified, (b) causally linked to the desired goals of quality education, (c) under the individual's control, and (d) a viable basis for rewarding effort and ability.

---

[15] Iwanicki, E.F. "Development and Validation of the Teacher Evaluation Needs Identification Survey". Educational and Psychological Measurement Vol 42, no. 1 (Spring 1982), pp. 265-74.

2.5 LINK TO ADMINISTRATOR APPRAISAL AND EVALUATION

In comparison with the volume of research and commentary on teacher performance appraisal, the educational literature on administrator appraisal and evaluation is fairly limited. While informal or ad hoc systems of administrator appraisal have been used in practice for some time, it is only recently that efforts have been initiated to examine this area more systematically.[16] To date, current practice (particularly among state-mandated programs in the U.S.) offers more specific and extensive information on the design and conduct of administrator appraisal than that which appears in the educational literature.

The issue of administrator performance appraisal is linked to teacher appraisal for several reasons:

(1) it is difficult to separate the effects of teacher behaviour and administrator behaviour on pupil progress and improvements in the quality of education;

(2) an effective system of appraisal encourages hierarchical co-ordination and accountability among all levels of the school organization;

(3) one of the criteria of administrator appraisal should be the administrator's willingness and ability to appraise teacher effectiveness; and

(4) administrator appraisal and evaluation promotes a better understanding among administrative personnel of teachers undergoing assessment.

The effective design and implementation of administrator appraisal systems is contingent upon an identification of the same appraisal issues (philosophy, objectives, type of criteria, and so on) as those outlined above for teacher appraisal. In addition, administrator appraisal programs confront many of the same challenges intrinsic to teacher appraisal, particularly:

---

[16] See, for example: De Bevoise, W. "Synthesis of Research on the Principal as Instructional Leader" Educational Leadership Vol 41, no. 5 (February 1984), pp. 15-20; Duhamel, R.; Cyze, M.; Lamacraft, G.; and Rutherford, C. "The Evaluation of Principals". Education Canada Vol 21, no. 2 (Summer 1981), pp. 20-27; Herman, J.J. "An Up-Front Approach to Administrative Evaluation". NASSP Bulletin (Nov. 1978), pp. 34-42; Lewis, A.C. Evaluating Educational Personnel op. cit.; McCleary, L. "Evaluation of Principals". Theory Into Practice (Feb. 1979), pp. 45-49; Renihan, F.I. "The Evaluation of Principals". Saskatchewan Education Administrator (Dec. 1980), pp. 22-31; Stow, S.B., and Manatt, R.P. "Administrator Evaluation Tailored to Your District or Independent School". Educational Leadership Vol 39, no. 5 (Feb. 1982), pp. 353-357.

(1) the inability of supervisors to observe behaviour on a day-to-day basis;

(2) the ambiguity of criteria for use in assessment of performance;

(3) the need to accommodate all personnel who have the right to a voice in determining the goals, process, and content of appraisal;

(4) the need to determine the relative contribution and relevance of alternative data sources (e.g., supervisors, self, subordinates, parents);

(5) the problem of reconciling the developmental and judgmental purposes of appraisal; and

(6) the difficulty of distinguishing the effects of uncontrollable variables in the work context from the effects of individual behaviour in assessing performance competency or goal achievement.

The essential elements in appraisal of principals include curriculum development, supervision of personnel, school management, and community relations.[17] School administrators have responsibilities which involve the co-ordination of expectations from politicians, parents, teachers, students, superintendents, and community groups. Flynn (1978), for example, states that "administrators who hold positions in school systems have regarded themselves as being in the centre of an array of demands".[18] Therefore, while one of the objectives of performance appraisal may be to establish consistent and reliable standards for use across different individuals performing similar tasks, some flexibility must be retained to allow each school administrator to respond to the unique circumstances and requirements of his or her particular environment.

There are currently very few validated instruments available for the appraisal of administrative performance. In a review of formal programs for measuring administrative competencies, Walters (1980) compared the four most well-known systems: Educational Leadership Appraisal (ELA), Georgia Principal Assessment System (GPAS), Individual Learning Materials (ILM), and Special Education Supervisor Training (SEST). Walters points out that "only GPAS attaches high importance to appraisal by persons other than the individuals themselves."[19] The primary use of most instrument systems is self-assessment and personal development.

---

[17] Deal, T.E.; Dornbusch, S.M.; and Crawford, R.A. "Villains as Victims: Evaluating Principals". Phi Delta Kappan, Dec. 1979, pp. 237-74.

[18] Flynn, J.J. "Evaluation of School Administrators". Comment on Education (October 1978), p. 5.

[19] Walters D.L. "The Measurement of Administrative Competencies". Phi Delta Kappan, Feb. 1980, p. 424.

Although teacher appraisal and administrator appraisal are conceptually and practically related, the educational literature treats these two topics separately. Future research and practice will need to address the question: how should teacher appraisal be linked to appraisal of administrators? Separate appraisal systems are designed around the necessary differences in tasks, responsibilities, and criteria for assessment between teachers and administrators. A system of appraisal is focused on the potential congruencies and interrelationships between teachers and administrators with regard to the philosophy, objectives, and process of evaluation. Future investigation into the development of a system of appraisal is likely to encourage examination of some critical questions surrounding the evaluation of educational personnel. Appendix 2 summarizes the major questions which future research will need to address.

As research progresses in the field of administrative assessment we are likely to witness an increase in attention to the relationship between teacher and administrator appraisal and the potential for a synergistic system of appraisal that recognizes the interactive qualities of performance in the educational context. The links between teacher and administrator appraisal are primary indicators that student growth is a shared responsibility.

The state of current literature on teacher performance appraisal is less a cause for concern than for optimism. At present, there is considerably more opinion in the literature than validated research, and what research has been conducted, particularly on teaching effectiveness, is inconclusive. In addition, current controversies surrounding the objectives and process of appraisal remain unreconciled. Yet the recent growth in vigorous attention to the topic of appraisal bears witness to the earnest attempts of researchers and practitioners to illuminate all sides of the evaluation question. It is such efforts that generate options in the design of appraisal and allow the development of a flexible appraisal program which considers the needs of all participants to the process. More importantly, the lack of consensus on particular details of appraisal has not obstructed general agreement that the overall values of honesty and fairness form the substratum upon which effective appraisal is built.

2.6 COMPARISON AND CONCLUSIONS

The foregoing review of the literature is a synthesis of research and commentary from both the general management and educational literature on performance appraisal. Both sets of literature are very similar in identifying several demands or requirements which must be satisfied by an appraisal system if it is to be effectively utilized in an organization. The list of possible uses of an appraisal system is essentially the same for both a business and educational setting.

a) The appraisal system can be used to generate data for personnel administration (compensation, training, counselling, and staff planning) and for personnel management decisions (retention, promotion, transfer, discipline). Decisions for personnel administration require appraisal data which basically indicates whether a person's level of performance is sufficiently satisfactory to retain his or her current job (i.e., performance at least satisfies the minimum standards). Decisions for personnel management require data from appraisal which distinguish levels of performance from one person to another, so that equitable and valid standards are used for decisions on retention, promotion, or discipline.

b) The appraisal system can be used for supervision and motivation of personnel (probationary review, communication, negotiation, and clarification of assignments and responsibilities, goal setting, delegation, support, and correction). Supervision of personnel via the appraisal system is accomplished through the communication and discussion of work assignments formally in a post-appraisal discussion and informally throughout the year. If a subordinate knows that these assignments will be subsequently reviewed in the next appraisal session, the motivation to concentrate on these assignments should be enhanced. In addition, if the appraisal system provides an opportunity to participate in the setting of assignments so that the persons are working essentially toward their own goals, then motivation should be enhanced as well.

c) The appraisal system can be used for introduction, maintenance, and/or changes in operating methods (clarification of operating methods, monitoring of methods, identification of training needs to change methods). If the organization is convinced that a particular philosophy or particular methodology of professional service is superior or desirable, then the appraisal system performance criteria can be designed to elicit and reward this particular behaviour. For example, if an educational institution believed that highly theoretical academic research was the path to organization excellence, then the appraisal system would use publications in refereed journals as a performance criterion. If the institution believed that teaching quality was the path to excellence, then student ratings of teachers and courses would be a criterion. If an organization wished to change its philosophy or methodology of professional service, then it would be necessary to develop a new set of performance criteria in the appraisal system to elicit and reward the new type of professional behaviour.

d) The appraisal system can be used for achievement of organizational and personal goals by translation of organization goals and plans into a series of individual goals and plans at all levels of the hierarchy. This approach to appraisal is commonly used where an organization has developed a mission or strategic plan that requires co-ordinated effort toward the achievement of organization goals. Departmental and individual goals can be made up of subsets of the organization goals, or individual task assignments can be derived from the larger goals. In either case, if individuals are held accountable for performance which contributes to the achievement of organization goals, then the likelihood of organization goals being achieved is enhanced.

e) The appraisal system can be used to protect the rights of individuals and provide due process for personnel decisions (fair and open criteria for appraisal, equity of treatment, equal rights or avoidance of discrimination, rights to negotiate and appeal). This use of appraisal is based on assumptions of validity in the appraisal system and open communication of performance standards and criteria. Open communications ensure that everyone is fully aware of the relevant standards of performance and will not be subject to arbitrary or biased standards. Validity ensures that people are appraised on their actual level of performance rather than on imagined or prejudged levels of performance. Such a system would not allow inadequate performers to hide, but more importantly would give full rights to recognition for competent performers who might otherwise not be recognized. These features of an appraisal system would go a long way towards ensuring equal rights and avoidance of discrimination in an organization.

f) The appraisal system can be used for professional development of groups and individuals in the organization. For individuals, this might come from feedback on their performance and discussions on how to improve (or how to continue) in the future. These discussions might even generate a set of action plans or personal goals on which to be appraised in the next evaluation. Training or development needs for groups may not necessarily be identified from the performance appraisal, but progress of a group toward some development goals could be assessed by aggregating some of the appraisal information. Development of individuals can also emerge from some of the other uses of appraisal, for example, progress toward learning, the organization's preferred style of professional service, or the opportunity to participate in goal setting.

Of course, not every appraisal system will use all these potential applications. This will be dependent on the needs or priorities of a particular organization. Secondly, different purposes or priorities for the appraisal system may exist for different levels in the hierarchy or different types of work within the same organization. Once the priorities and purposes of appraisal for different parts of the organization have been determined, then attention can be drawn to the methods of appraisal. It is clear from the review of both sets of literature that method of appraisal is a complicated and frustrating topic because all methods involve various compromises of cost, convenience, and complexity. The requirements for good appraisal methods are generally well understood, but the requirements are so demanding as to be practically impossible to satisfy by any one method of appraisal. Consequently, we are led to a conclusion that an organization must consider using a multifaceted approach to appraisal methods.

While a multifaceted approach is also complicated, it is possible to reduce some of the complexity by using different methods of appraisal for different types of teachers or positions at different levels in the hierarchy of a school. By way of illustration, Exhibit 3 presents a hypothetical example of how appraisal methods might vary in one school for different types of teachers. Each of the various methods in a multifaceted system must satisfy the criteria of validity, reliability, and acceptability, but these criteria can actually be enhanced by using different appraisal methods for different types of work. Moreover, a multifaceted approach avoids the problem of burdening everyone in the hierarchy with having to go through the administrative bureaucracy of appraisal methods which are not appropriate for different types of work.

Exhibit 3

Multifaceted Appraisal System: A Hypothetical Example

Type of Teacher	Focus of Appraisal	Evaluation Criteria	Possible Methods of Appraisal
1. Entry level or probation	decisions on retention or training	skills, traits knowledge, potential	credentials; rating and behaviour scales; incidents; simulations; possible ranking
2. Experienced permanent staff	maintain standards, professional development	continued teaching effectiveness, non-classroom work, development needs	observations; peers; secondary indicators; activities; negotiated plans
3. Senior teaching staff	maintain standards, wider responsibility	continued teaching effectiveness, administrative or leadership behaviour, results of plans	observations; negotiated work planning and review; results-focused appraisals.
4. Special performance review	employee assistance, upgrading performance standards, dismissal	specific to problems, symptoms or indicators, specific timing and activity plans	negotiated plans, activities and targets; specific observations; trial periods; due process mechanisms.

For example, it may be futile to require a novice teacher to go through an elaborate goal-planning activity for appraisal when the primary concern is whether the teacher possesses minimum acceptable standards of knowledge and skills for teaching in a subject area. Similarly, for an experienced teacher with a long history of excellence in teaching, it may be futile to force the person and the principal to go through a charade of detailed goals-planning when the primary concern is to show recognition and appreciation for good work. On the other hand, for a senior teacher who wishes to qualify for higher levels of responsibility, a detailed career development plan would be quite appropriate both for professional development and for performance appraisal. Thus a contingency approach seems to be the most suitable for the question of choice of appraisal methods. This in turn requires that senior administrators in a school system have a systematic knowledge of strengths, weaknesses, and applicability of different appraisal methods.

Finally, another important consideration is the cost and complexity of administering an appraisal system. Many business organizations in the past have opted for simplistic methods of performance appraisal in order to go through the motions of appraisal as quickly and cheaply as possible. While all the faults of this approach are fully recognized, it is also recognized that virtually any system which tries to satisfy all the demands on a performance appraisal system will be very complex and cumbersome to develop and administer. This will be true for a multifaceted approach to appraisal as much as for any other modern approach to appraisal. However, a type of multifaceted approach can be used which can reduce some of the complexity and administrative burden of an evaluation system.

To reduce some of the complexity, it is necessary for an organization to define or establish certain minimum acceptable standards of performance (i.e., a respectable average level of performance) and develop some simple (valid, etc.) indicators of whether a person's performance is in this range. If the primary concern of the organization is to ensure that the great majority of professionals are performing at or better than the acceptable standard of performance, then no other form of appraisal need be called into play. This would be a maintenance approach to management, much like setting a ship on a course and then making only minor corrections to the steering to remain on course. Other, more complex forms of appraisal would be used for those persons whose performance is not yet up to, or has fallen below, standard, or for those persons who wish to go through professional development to a different area of performance. Thus an array of appraisal methods would be available for use in a school system, but a "Pareto's optimum" approach could be used for the 20% of staff requiring special appraisals and simpler appraisals could be used for the 80% of staff requiring routine appraisals.

So far, the conclusions presented have assumed commonalities between the general management and the teacher evaluation literature on performance appraisal. We now turn to some of the major differences between the two sets of literature.

a) Career progression vs. organization tenure - While employees in a business organization must perform up to some level of acceptable standards to retain their jobs, there is a strong emphasis in such organizations on being able to qualify for higher level positions. The culture of expectations in business organizations is that one must expect to have vertical career progression in the organization hierarchy. Thus the appraisal system must have a dual focus of giving feedback on current or past performance and also making assessments of potential for different kinds of work in the future.

In a public school organization, there is far less expectation of vertical career progression and instead the emphasis is on maintenance of organizational tenure or a culture of continuing to perform professional service work for the duration of a career. This is typical of professional service bureaucracies and the appraisal system consequently focuses essentially on the performance of professional service work, without any particular reference to potential for different kinds of work. Some professional bureaucracies do provide dual career ladders to a certain extent, to allow for career progression within professional service work without a person having to join the administrative hierarchy for career progression. However, even in these cases, the emphasis in performance appraisal continues to be on how well the person performs in the professional service tasks.

b) Output vs. process appraisals - Business organizations tend to have a greater use of output criteria in appraisal systems for a number of reasons. One is that many businesses have the opportunity to count or measure outputs in a variety of ways due to the nature of the production process in many businesses. Secondly, most businesses use some form of management by objectives (MBO) to manage business operations and it is a natural extension of this to use an MBO form of performance appraisal. Thirdly, there is less concern with how results are achieved (save for legal and ethical constraints) than for what is achieved in business organizations. This may be due to lack of confidence in any one style or process for best obtaining results, or due to a belief that many alternative processes are possible for achieving the same ends. In any case, the preference is to focus appraisals on results more than process.

In educational organizations, the dominant focus of appraisal systems is far more on process than on outputs. This too is consistent with professional bureaucracy theory[20] but also may be due to the difficulty (or reluctance) of measuring or counting outputs. Perhaps this is a function of public schools being indirectly financed by public tax resources, as privately financed schools seem to be much more willing to use output-based performance appraisals. The widespread use of process-based appraisals is also interesting in the light of Levin's[21] comment that so little satisfactory knowledge exists about which teaching methods do the best job of education.

c) Performance-based vs. seniority-based compensation - Consistent with use of output-focused appraisals is the use of appraisals for determining pay differentials according to performance differentials. Most business compensation systems provide some universal cost-of-living pay adjustment, but rate of progress through a job salary range or progress to a higher salary range is determined by assessment of performance. Consequently, the appraisal system in some way must be concerned with ranking or comparing one individual's work to another. Similarly, business organizations are far more concerned with justifying compensation dollars and therefore are far more prone to dismiss unsatisfactory performers. This too causes the appraisal system to have to rank-order performance evaluations in some manner.

---

[20] See, for example, Mintzberg, H. The Structuring of Organizations. Englewood Cliffs, N.J.: Prentice-Hall, 1979.

[21] Levin, B. "Teacher Evaluation: A Review of Research". Educational Leadership (December 1979), pp. 240-245.

In school organizations, the emphasis seems to be on acquiring entitlement to compensation adjustments by virtue of seniority and maintenance of adequate performance. Pay raises tend to be uniform with no performance differentials and progression through a salary range is by accumulation of years of seniority. Consequently there is no need for the performance appraisal to rank-order performance or otherwise distinguish between levels of performance except to identify serious cases of below-standard performance.

d) Non-union vs. unionized personnel - Few people at the professional service level in business organizations belong to bargaining units although they may belong to professional service associations. Consequently, their employment contracts or conditions of employment are based on uniform company personnel policies and standards plus any individual aspects of employment conditions which they may be able to negotiate. Performance appraisal will follow a similar pattern: a universal floor of common standards and practices for all, plus individuated aspects of appraisal based on the person's unique circumstances and role in the organization. In practice, there is not a lot of variation in the overall performance appraisal system used for different individuals, but the principle of individuation exists for non-union personnel.

In teaching organizations, virtually all certificated personnel are members of bargaining units and a basic union philosophy is equal rights and treatment for all members. Consequently, the nature of the appraisal system tends to be universalistic as well. This is consistent with the approach to compensation discussed in section (c) above. Individuation in teacher evaluation does exist, of course, depending upon the strengths and weaknesses of individuals and their relationships with principals, but the influence of the union is to try to maintain universal equivalence as much as possible.

e) Personal vs. non-personal accountability for results - In business organizations, an operative assumption in performance appraisal is that individuals are essentially responsible for the achievement of assignments or agreed-upon results. They are given a certain amount of resources and a certain amount of time and are expected to be the catalyst which produces the outcomes. Many results-focused systems have provision for extenuating circumstances, unforeseen events, and so on, but ultimately the individual must accept personal accountability for the results produced from their domain of responsibility.

In contrast, it has been strongly argued in the literature on teacher evaluation that individuals cannot be held fully accountable for results in the public education system because of too many uncontrollable factors or extenuating circumstances. Instead, the teacher is seen as just one of a series of actors in an educational environment which ultimately produces some form of education. Without discussing the merits of this view of accountability, it is quite clear that the viewpoint on personal accountability for results is quite different than that found in business organizations.

f) Internal vs. public accountability - In business organizations, the conventional wisdom is that boards of directors or senior executives accept the public accountability for the performance of their organization in society. Rarely, if ever, does the public try to be involved in the internal operations of the organization except through communication with senior management. This means, in general, that accountability for individual performance within the organization is an internal matter between supervisors and subordinates. The work of individual professional service workers is less subject to public scrutiny and if the public does have some concern about the quality of this work, the matter is taken up with management rather than with the professional (an exception would be a direct fee-for-service relationship, such as with a doctor or lawyer).

In public schools, there is far greater public scrutiny of the works of individual teachers and the public is equally likely to confront the individual professional or the management of the organization. This is true even though there is no direct fee-for-service relationship between the public and the professional. The impact of this difference in concepts of accountability is that teacher appraisals have to be more sensitive to public accountability than would be true in most business organizations. This raises an aspect of complexity in teacher evaluations which would not be generally found in business organizations.

g) Whole vs. narrow performance evaluations - A "whole" job orientation in performance evaluation refers to the concept of evaluation on the complexity or multidimensional aspects of a job. If a job is relatively rich in content (i.e., extensive decision-making, autonomy, responsibility, multiskilled, long time horizons for results), then the appraisal system ought to be sufficiently multidimensional as well to be able to reflect or capture the richness of the job. If the job is relatively low in complexity or essentially unidimensional, then the appraisal system can be correspondingly unidimensional or focus on a narrow, but important aspect of the job. Despite the tendency in many business organizations to try to find a relatively simple and painless appraisal system, there is little tendency to try to focus an appraisal system on just one aspect of work. Business organizations seem to recognize that successful performance is a multidimensional phenomenon and there is an ongoing search for appraisal systems which can adequately reflect that complexity. In educational organizations, there is also a recognition of the complexity of a teacher's job, but the common practice is to focus appraisals more narrowly on specific classroom performance. Thus appraisals in schools tend to be more narrow in orientation and appraisals in businesses tend to be more whole-job oriented.

This situation does not pose a problem if the narrowly focused appraisal is concerned with a dimension of performance that is crucial to the organization or is a good indirect indicator or surrogate of the balance of the professional's work. For example, one might say that there is more to being a teacher than teaching in the classroom. However, if classroom performance is a valid representation of a person's overall performance in the school, then it would be sufficient to rely on the relatively narrow focus of evaluating certain classroom behaviours to make inferences about other dimensions of a teacher's performance (i.e., good classroom behaviour indicates good relations with parents, indicates good planning of lessons, indicates good grading of assignments).

If these assumptions are not valid and the narrow aspect of a job which is the focus of evaluation is not a good indicator of other important dimensions of performance, then the appraisal system may be of dubious validity. For example, if classroom behaviour is independent of other important teacher behaviours, then a narrowly focused appraisal system does not really reflect the wholeness of a person's work. This situation would be exacerbated if a narrowly focused appraisal system unfortunately evaluated performance on some relatively unimportant dimension of work (i.e., neatness, punctuality) that was easy to measure, but of inconsequential validity. The question then for narrowly-focused teacher evaluations is twofold: which are the key dimensions of performance on which evaluations should focus, and are these dimensions adequate indicators of other important aspects of a teacher's work? The alternative is to develop multidimensional appraisal systems similar to business organizations which try to capture the complexity of the work of a professional.

In summary, all of the above points suggest a series of cultural or philosophical differences which exist between performance appraisal in business and teacher evaluation. These are differences which have evolved over time and reflect the different nature of operations in the two types of organizations. However, the differences are not immutable and some possibilities for convergence of practices can exist. These convergences should not be sought merely for the sake of convergence. In keeping with the theme of conclusions from this report, appraisal practices from the business world should only be borrowed when they make sense to the needs and priorities of a particular school organization or for a particular segment of the certificated staff. Similarly, historical precedents in teacher evaluation should only be retained or perpetuated if they continue to make sense or meet the needs of a particular school system. Ultimately such judgements will have to be made by persons involved in each particular school system.

3. FIELD SITE RESULTS

3.1 INTRODUCTION

The purpose of this project was to conduct a series of field study site visits to organizations which are analogous to public education organizations in that large numbers of professionals are employed and formal appraisals are conducted of professional service performance. The field site study was organized into two parts: a study of ten analogous organizations in Ontario which could provide useful insights into the appraisal of professionals, and a study of selected public school boards in the United States which have acquired a reputation for being very progressive in their approach to appraisals of educational staff.

In general, the study of analogous organizations focused upon the following issues:

i) to identify the appraisal methodologies and the uses of appraisal in the various types of organizations;
ii) to identify major differences (if any) in approaches to appraisal which seem to be required by the different organizational types;
iii) to identify differences in approaches to appraisal in different functions of the organization and in different levels of the hierarchy;
iv) to identify whether process or output approaches to appraisals are typically used in some functions and organizations and not in others;
v) to identify various approaches to developing criteria and gathering data on the work of professionals; and
vi) to identify linkages between the appraisal process and other parts of human resource management in the organization, such as compensation systems, training and development, career planning, and supervision.

3.1.1 The Business Group

The Business group was identified by consultation with colleagues at the Personnel Association of Ontario and the Compensation Research Centre of the Conference Board in Canada. The criteria for sample selection were employment of a high proportion of professional service employees and a good reputation in the personnel community for general personnel management. From this it was assumed that such organizations would also be examples of good practices in performance appraisal since performance appraisal is a foundation for so many other aspects of personnel management. The identity of each organization has been kept confidential, partly to encourage free-flowing, frank discussions about performance appraisal and also because some sites requested confidentiality as a condition for participating in the study. Out of a list of ten recommended sites, five organizations agreed to participate when first approached and the remainder of the list was not used.

3.1.2 The Colleges Group

Colleges of Applied Arts and Technology (CAATS) were selected for this study because they obviously employ a large number of teaching professionals and are located in the Ontario educational environment. They are different from schools in public education in that they are post-secondary, tuition-based schools and many of their instructors are not formally trained or educated as teachers. However, the colleges can be a useful sample to study for performance appraisal practices because they have to be more directly accountable to their students for a good education (tuition fees) and to the local employment community. Many colleges use advisory boards of local employers and prepare students for Provincial trade certificates or licences. Consequently, the performance of a college in producing qualified students can be closely scrutinized in many ways. It follows from this that many of the colleges have devoted careful attention to maintenance and upgrading of teaching quality.

The colleges are required by their collective agreement with teachers to provide appraisals of teaching effectiveness once a year. The method of appraisal and the dimensions of performance are not specified and it is up to each individual college to devise its own system. Practices ranged from lip service to appraisal, to different approaches within different departments in the same college, to integrated systems with quite serious attention paid to appraisals.

The sample for this study was identified by a study group of CAAT personnel directors for the purpose of sharing information on performance appraisal practices at each college with the eventual hope of developing some common (voluntary) standards across the Province. The study group of personnel directors was able to identify five colleges as being the most advanced or at least the most seriously concerned with developing effective appraisal systems. These five colleges are identical in types of professionals employed (lecturers, teaching masters, area co-ordinators, and department heads). Although these colleges were not selected to be a random or representative sample, there is no particular geographic region, program specialty, or size factor in common among the group studied, so they represent a reasonably good cross-section of the colleges in the Ontario System.

3.1.3 The U.S. Schools Group

The third type of field site sample was a group of United States public school systems which had been identified through the literature review as being front-runners in either innovative or serious approaches to teacher evaluation. Although the U.S. Schools group obviously operate in different education and employment environments than Ontario, they do offer the opportunity to examine a number of state-of-the-art approaches to teacher evaluation which might be applicable or relevant to Ontario schools. Secondly, many of the States have state-wide, mandatory teacher evaluation systems which serve as a baseline for evaluation practices in local school districts.

The U.S. Schools group consisted of seven public school systems as well as the teacher evaluation service of Educational Testing Services (ETS) of Princeton, New Jersey. The seven school systems are:

i) The Chicago Board of Education, Chicago, Illinois;
ii) Montgomery County Board of Education, Rockville, Maryland;
iii) High Point Board of Education, High Point, North Carolina;
iv) State of Connecticut Department of Education, Hartford, Connecticut;
v) State of Georgia Department of Education, Atlanta, Georgia;
vi) Oakland Unified School District, Oakland, California;
vii) Orange County Board of Education, Orlando, Florida.

There is no particular geographic or urban pattern to the U.S. Schools group. Sites range from large cities (Chicago, Oakland) to smaller cities (High Point, Orlando), from northern industrial states (Connecticut, Maryland) to Southern (Georgia, Florida). It was hoped that such a distribution of sample sites would provide a diverse range of approaches and concerns in teacher evaluation systems and not be limited to the special problems of any one geographic or socio-economic area.

3.2 PROCEDURE

Each of the prospective Business and College sites in Ontario was contacted by telephone to arrange interviews with the director of personnel or the chief administrator of the performance evaluation system. Subsequent interviews were then arranged with department heads or supervisors who actually conducted appraisals to obtain their perspectives on their experiences with the system.

Occasional interviews were also possible with professionals themselves on their experiences with being appraised by these systems. No attempt was made to try to gather widespread evidence from appraisees on satisfaction with a system. This would have required extensive surveys with numerous samples of respondents and the logistics of trying to conduct a valid and reliable study of that sort was beyond the resources of the field study project. Instead, satisfaction with a system had to be inferred from three admittedly weak indicators: testimony on satisfaction from the interviews, numbers of revisions felt to be necessary to a system, and record of grievances, appeals, arbitrations, and lawsuits arising from appraisals.

Site interviews with the Business and Colleges group were conducted during a period from August to December, 1982. Consequently, the information gathered reflects practices and systems in effect or planned at that point in time. Prospective sites in the U.S. Schools group were contacted by telephone by one of the research team to arrange a site interview. All sites approached were willing to participate in the study and site visits were done during the period September 1982 to May 1983. Interviews were conducted with the senior administrator responsible for the teaching evaluation system and occasionally with specialists responsible for design, implementation, monitoring, or evaluation of the system. Interviews at all school sites were also conducted with supervisors (superintendents, inspectors, principals) who used the appraisal system and, occasionally, interviews were done with teachers subject to the evaluation system. The same indicators for relative satisfaction with a system were used for the U.S. Schools sites as were used for the Business and Colleges groups.

3.3 RESULTS

Results will be reported for the following topic areas:

a) Initial Design and Implementation of the System;
b) Current Objectives and Purposes of Appraisal;
c) Appraisal Design and Techniques;
d) Criteria or Standards for Appraisal;
e) Sources of Data on Performance;
f) Training and Preparation for Apraisal;
g) Special Cases of Appraisal and Organizational Justice;
h) Evaluation and Maintenance of the Appraisal System;
i) Trends and Outlook for the Appraisal Systems.

3.3.1 Initial Design and Implementation of the System

A comparison between the three types of samples can be made by looking at reasons for developing the appraisal system and the method by which the system was designed and implemented. Table 1 summarizes these comparisons. As can be seen from this table, there is a fairly clear distinction between the Business group and the two groups of educational organizations in the reasons for developing an appraisal system. The Business group reasons, in order of importance, were (1) improved organization performance, (2) improved personnel administration, and (3) improved quality of professional service. In both types of educational organizations, the reasons were (1) mandatory requirements for teacher appraisal, (2) improved quality of professional service, and (3) improved personnel administration. These priorities may differ in individual organizations but the order of priorities represents the patterns for the sample as a whole.

In terms of design and implementation of a system, the top management approach to design is predominant in the Business group while the joint or collaborative approach is favoured in the educational organizations. It is possible that a relationship exists between reasons for developing a system and method of developing a system. The business organizations were primarily concerned about improving organization performance and used a top management approach to this end. The educational organizations had a high priority on quality of professional service and used a collaborative approach to the design of an appraisal system. However, this pattern is less clear for the Schools group since three of them used a top management approach and were also concerned about quality of education. Two of the schools with the top management approach (Chicago, Oakland) were in a politically demanding situation and may have behaved similarly to the Business group under pressure to produce system-wide improvements in a short period of time.

Table 1

Initial Design and Implementation of the Appraisal System

	Business Group (5)	Colleges Group (5)	U.S. Schools Group (7)
	(order of importance)		
A. Impetus priorities for developing a new system			
1. Improve organization performance	1	-	3
2. Mandatory requirement	-	1	1
3. Improve quality of professional service	3	2	2
4. Improve personnel administration	2	3	-
	(frequencies)		
B. Methods of development of the appraisal system			
1. Top management design and development	4	2	3
2. Joint committee design and development	1	3	4

3.3.2 Current Objectives and Purposes of Appraisal

Table 2 provides a summary comparison between the three samples on a number of general purposes served by the appraisal systems. Some of the purposes may be implicit in a number of organizations (i.e., to improve student learning in schools), but are only listed in this table if they were explicitly articulated at the sites. An example of this concept is the use of appraisal information for summative evaluation and personnel decisions, such as compensation, tenure, and promotions. The majority of the Business group are explicit in stating that such information has a direct linkage to personnel decisions, but the majority of schools and colleges refer to such information as "possible inputs" to personnel decisions. This means that appraisal information may be called into play if performance problems arise or are identified in some way, or that personnel decisions such as compensation progression and seniority accumulation occur automatically as long as a person's performance is "satisfactory". However, all of the sites studied admitted that appraisal information has to be the basis for taking action and justifying such actions if performance problems occur. In these situations there is a direct link between formative performance appraisal and summative personnel decisions in all organizations.

Table 3 provides a summary comparison of more specific, routine purposes of appraisal systems between the three samples. The major differences appear in aspects which are more characteristic of business organizations (i.e., performance pay, promotion potential). Another major difference is that identification of training needs is a common purpose of appraisals in business organizations. The purpose should be just as prevalent in the U.S. Schools group given the emphasis on professional development found in most of these sites and the absence of this purpose suggests that the emphasis on professional development in the Schools group is not consistently followed through for individual teachers. Instead, the general approach to professional development is feedback, coaching, and guidance; general school board professional development activities; and individual initiatives in training activities.

3.3.3 Appraisal Design and Techniques

Table 4 provides a summary of the information in the three samples. From Table 4 it can be seen that all three types of samples have a hierarchical difference in approach to appraisals. Without exception, administrators and managers use a goal-setting and action-planning style of management and their performance is evaluated either on the completion of their plans or the results they achieve. In general, this goal setting and planning by administrators is done in the context of organization goals and plans. Between divisions or departments, a uniform format of appraisals is used by sites in both the Business and U.S. Schools groups. Only in the colleges were some examples found of different appraisal practices being used for jobs at an equivalent level in different departments in the same organization. While it can be argued that the nature of work may be sufficiently different between departments to justify different appraisal methods, it should be recognized that such practices are open to claims of inequity in the organization. The other difference in general is that most of the U.S.

Table 2

Current Objectives and Purposes of Appraisal Identified for the Organizations in the Study

Objectives/purposes	Business(5)	Colleges(5)	U.S. Schools(7)
	(number who claim these objectives)		
1. Improve student/client learning	1	1	1
2. Improve quality of service/teaching	2	5	3
3. Improve organization management	5	2	2
4. Improve professional development	5	3	4
5. Provide a direct link to personnel decisions	4	-	-
6. Provide inputs when required for personnel decisions	1	5	3

Table 3

Specific Purposes of Appraisals

Specific Purposes	Business Group (5)	College Group (5)	U.S. Schools (7)
	(number who	claim these	purposes)
1. Seniority maintenance	3	5	7
2. Eligibility for performance pay	3	0	0
3. Identify training needs	5	5	0
4. Discover performance problems	1	0	0
5. Document performance problems	5	5	7
6. Assess promotion potential	5*	0	0

*Note: The Business group distinguish between assessing competence as a professional and assessing potential for promotion. A good record of performance as a professional is a prerequisite to eligibility for promotion. Special performance appraisal techniques (i.e., assessment centres, trial runs as supervisors, acting project coordinator roles, etc.) are used for appraisal of promotion potential.

Table 4
Appraisal Design and Techniques

Types of Design Techniques	Businesses (5)	Colleges (5)	U.S. Schools (7)
a) 1. Differences in hierarchy	5	5	5
2. Differences across divisions	1	3	0
b) 1. Focus on certification/competency	1	5	6
2. Focus on work processes	1	5	7
c) 1. Focus on action plans	4	4	3
2. Focus on results	4	1	3
d) Formats of appraisals			
1. Summary essays, memos, letters, etc.	3	5	3
2. Tests, rating scales	2	1	4
3. Student/client evaluations	1	5	2
4. Checklists of abilities, skills	4	4	3
5. Forced-choice comparisons	0	0	0
6. Critical incidents record	1	0	1
7. Ranking	2	0	0
8. Semi-structured observations	0	3	6
9. Unstructured observations	0	0	1
10. BARS, BORS	0	0	0
11. MBO - action plans	4	4	3
12. MBO - results	4	1	3
13. Self-evaluations	2	3	3
14. Peer evaluations	1	0	0
15. Numerical summaries of performance	3	0	2
e) Interim feedback	5	5	4

Schools had a strong emphasis on certification or competency evaluation for entry level jobs which was not found in either the colleges or the business organizations (with the exception of one accounting firm). This means that many of the U.S. Schools group had to verify competency after a person was employed while the other organizations assumed competency at the time of hiring and instead concentrated on verifying level of performance.

In terms of the focus of appraisals, both the colleges and the U.S. Schools had a pronounced focus on process in their evaluation systems, while the Business group had a pronounced focus on results. Some of the U.S. Schools had a hybrid approach by focusing on process and results but this was less frequent in the other two samples.

3.3.4 Criteria or Standards for Appraisal

Table 5 summarizes the information in this section for the three samples of organizations. There is a roughly equal tendency among all three groups to have general job descriptions for professionals and little tendency to have very detailed job descriptions. However, two of the U.S. Schools do not have job descriptions at all for teachers which is a significant difference from the Business and Colleges groups.

All three samples have established licence, degree, or certificate requirements as minimum basic standards for qualifying for a position but only some of the U.S. Schools group do their own certificate testing. The practice of defining specific competencies and professional service behaviours is far less common in the Business group than in either the College or U.S. Schools groups. The specific competencies and behaviours in the colleges are defined in the student evaluation forms and in the U.S. schools by competency criteria.

It is fairly common in all three samples to have some specification of professional service behaviours or styles as part of the appraisal criteria. In the Business group, these were agreed upon methods or corporate styles for professional service work. In the colleges and schools, these were criteria specifications for the competencies required. (The criteria specifications are often written in behavioural terms or guidelines.) Only one of the Business group (accounting) also used specific process standards and criteria for appraisal, while these were far more commonly used in the colleges and schools. Typical criteria consisted of teachers' lesson plans and teaching activities which then formed the basis for observations by principals, or other supervisory personnel. This approach of using rating scale or semi-structured observations of a professional's work by a supervisor was practically non-existent in the Business group.

Instead, the Business group was consistent in the use of results specifications for appraisal, both in terms of action plans and results specifications of outcomes. The colleges and some of the schools were active in specifying plans for objectives, but rarely were results themselves specified. This means that completion of the plans themselves were accepted as results for evaluation purposes rather than verification of whether the plans actually achieved the intended objectives.

Table 5

Comparison of Standards and Criteria of Performance

Types of Standards and Criteria	Business (5)	Colleges (5)	U.S. Schools (7)
a) Job descriptions - general	4	5	3
- detailed	1	0	2
b) Minimum standards specifications			
- licence, degree or certificate only	5	5	7
- specific competencies; behaviours	2	5	4
c) Process standards specification			
- professional service behaviours, styles	4	3	6
- specific process plans, activities	1	4	7
d) Results specifications - plans	5	4	4
- outcomes	5	1	1
e) Professional development - plans	5	5	5
- outcomes	3	0	1
f) Contribution to organization plans	5	3	2
g) Compatibility with organization culture			
- explicit	3	2	1
- implicit	2	0	3
h) Appraisal orientation - whole	5	4	2
- narrow	0	1	5

Professional development plans were used in all of the businesses and colleges and in many of the schools. The businesses tended to look more closely at the results of the development plans and therefore these were used more seriously for appraisal purposes than in either colleges or schools. In all cases, development plans and results were examined very carefully in special review situations of problem cases in performance. Another pronounced difference between performance criteria in business and the two types of educational organizations was the emphasis in the Business group on contribution to organization plans and compatibility with organization culture. These criteria are consistent with the greater use of goals-based management and appraisal systems in the Business group, and the effect is to have a wider or more of a whole-job approach to appraisals in the Business group. Despite the widespread use of MBO at the administrative levels of both the colleges and schools, instructors and teachers are evaluated by a much narrower range of criteria and consequently there is a narrower operational concept of the work of the professional.

3.3.5 Sources of Data on Performance

Table 6 summarizes the results for each of the three groups of organizations. The Business and College groups have much more frequent formal appraisals than the U.S. Schools although the average practice across all three samples is to have a formal annual appraisal. The use of bi-annual and tri-annual formal appraisals in the U.S. Schools sites was generally reserved for senior, tenured teachers in whom principals had a lot of confidence. Provision was made in all organizations for shorter term formal appraisals where problems seemed to exist. However, the schools with up to three years between formal appraisals do run a greater risk of performance getting out of hand or being difficult to correct, by the time that it is concluded from formal appraisals that a performance problem exists.

Despite frequent recommendations in the literature, the use of formal pre-observation conferences was not common in either the colleges or the schools. Many of the educational sites had such conferences as an optional practice, but the exercise of the option was not especially frequent. This seems to be a shortcoming in appraisal practices where so much depends on classroom observations by principals or department heads. As was stated previously, many supervisors follow standard observation forms or rating scales in conducting the observations so perhaps it can be argued that there is not much to be discussed in a pre-observation session. It was also claimed at some sites that the sheer magnitude of the number of required classroom visits prohibited much opportunity for discussion beforehand (and afterwards). However, for those systems which require their teachers to develop teaching strategies, lesson plans, and so on, it would seem necessary for a principal to have a pre-observation conference in order to establish what the principal was supposed to be looking for in the classroom visit.

The Business group do not use any form of formal pre-observation conferences but instead rely on the annual planning and goal setting discussions to accomplish the same purpose. These discussions also alert the supervisor as to what to expect and look for in a subordinate's performance, but the time period could be for up to a year afterwards and the observations less formal.

Table 6

Sources of Performance Appraisal Information

Sources		Business (5)	Colleges (5)	U.S. Schools (7)
a) Formal Appraisals	i) semi-annual	4	2	-
	ii) annual	1	3	4
	iii) two years	-	-	1
	iv) three years	-	-	2
b) Formal pre-observation conference		0	2(+3 optional)	2(+2 optional)
c) Formal post-observation conference		5	5	7
d) Formal direct observation	- supervisors	1	5	7
	- experts	0	0	1
e) Formal self-reports of performance		3	3	2
f) Formal reports by clients/students		2	5	2
g) Formal reports by subordinates		0	2	1
h) Formal measures of client/student performance		2	2	2
i) Supervisor monitoring and informal observations by supervisor		5	5	7

Post-observation conferences are used by all of the organizations in the study. The difference between the educational organizations and businesses is that the conferences in colleges and schools take place very soon after the formal observation of performance. In the Business group, the post-observation conference is the semi-annual or annual appraisal conference which serves as the summative evaluation of the overall performance for the period. In terms of learning theory, the practice in educational organizations is more effective because feedback occurs shortly after the observed performance. The effect is similar to a formative appraisal even though the intent is to be summative. On the other hand, the practice in business is to try to take a whole job perspective for the whole period of performance and these appraisals would not be as susceptible to recency bias as might a teacher appraisal.

Formal, direct observation by supervisors was the most common source of appraisal information in both of the educational groups and was almost nonexistent in the Business group. This means that educational organizations rely extensively on observations by department heads or principals as the most basic source of appraisal information. Although they are aided by checklists, rating scales, criteria references, and observation guides, the performance appraisal of a teacher ultimately boils down to what a supervisor observes in a classroom. Obviously this places a considerable burden on such supervisors to be competent observers so that the data derived from this source are valid and reliable. The other obvious problem with formal direct observations by supervisors is the short time period, the relative artificiality, and the narrow focus of such an approach.

To compensate for these limitations, most organizations tend to use some supplementary sources of information about a person's performance. A few in each sample used self-reports on performance by subordinates/teachers, and some used reports or evaluations by clients or students. Student evaluations of courses and teachers were used in all of the colleges, due to the high emphasis by CAATs on good quality teaching (and the implicit assumption that students could provide valid information about teacher performance). The colleges were also careful to try to distinguish between evaluation of a course (which is akin to program evaluation) and evaluation of an instructor. Formal reports or ratings by subordinates and teachers were used in a few cases in colleges and schools for the appraisal of principals or department heads. There did not seem to be any problems with acceptance of this procedure.

Formal measures of client/student performance were used in a few sites in each of the three types of organizations. In all cases, this information was used more for program evaluation or evaluation of an organization than for individual professionals within the organization. Of course program and organization evaluation can be used for evaluation of the responsible administrator because it is consistent with the use of a results-based appraisal system for administrators and managers. No evidence was found of student performance criteria being used for the evaluation of individual teachers or instructors, although this could be done if some individual professional were found to be responsible for the negative (or positive) results of such program evaluations.

Finally, the most common source of appraisal information in all three samples was monitoring and informal observations by supervisors. As stated previously, this sort of information is inevitably acquired as a result of the routine interactions between people working together in an organization. It is part of the supervisor's role in an organization to convey expectations about performance, to give feedback on whether performance is consistent with expectations, and to provide help and counsel where required, or support and recognition where merited. In the course of these duties and responsibilities, the supervisor cannot help acquiring a cumulative series of impressionistic judgements about a subordinate's performance. In addition, supervisors typically find themselves to be recipients of a variety of unsolicited information about performance in the organization. Unfortunately these are often complaints, but occasionally can be compliments. As well, if there happens to be a sufficiently well established culture peculiar to the organization, supervisors can find themselves in the unofficial (and unconscious) role of socialization agent for the organization.

All of these factors combine to place a burden of responsibility on supervisors similar to that described for direct observations of performance. However, the supervisor is not usually aided by a clear set of guidelines, competencies, and criteria for making informal observations and impressionistic judgements. Supervisory monitoring and informal observations provide a rich source of data about a person's work performance and help to capture the wholeness of a job better than many of the formal sources of data. The challenge is to minimize the likelihood of bias or limited validity and reliability in such data. The existence of this challenge highlights the need for training in performance evaluation for supervisors since they seem to be the key link in virtually any form of performance appraisal found in this study.

3.3.6 Training and Preparation for Appraisal

Table 7 provides the summary results for this section. The results previously reported in this study have established that the supervisor plays a pivotal role in evaluation of performance. The supervisor is the primary communicator/negotiator of expected standards of performance, the primary observer or data gatherer about performance, and finally, the primary deliverer of feedback to subordinates about how their performance is perceived and evaluated. In recognition of the importance of these supervisory roles, it should follow that organizations provide considerable support and training for supervisors on how to do performance evaluation.

All of the sites in the three samples in the study had extensive manuals, procedures, and policy guidelines on performance evaluation available for supervisors. Support was also provided by the respective personnel departments in the form of administration of the system to ensure that appraisals were done on time, and according to procedures. (Some might perceive this as pestering by the personnel department rather than support.) Direct support was also provided in most business, colleges, and schools by having supervisors discuss

Table 7

Training and Preparation for Appraisals

Types of Training and Preparation	Business Group (5)	College Group (5)	U.S. Schools Group (7)
1. Formal support systems	5	5	7
2. Informal support systems	5	5	7
3. Formal training for supervisors	2	1	2 (+2 optional)
4. Informal learning for supervisors	5	5	7
5. Formal Orientation for subordinates	5	5	7
6. Informal learning for subordinates	5	5	7

Note: Organizations could obviously use formal and informal practices together.

appraisals with someone from a more senior level of management. This secondary level of discussion accomplished two purposes. Senior managers would be kept in touch with general levels of performance of a wide range of personnel and evaluations of personnel between divisions or departments could be monitored. This procedure was important for the early identification of performance problems and for ensuring equity of treatment across divisions in a particular organization.

Formal training in conducting performance appraisals was quite different between the samples. Only two of the Business group and one of the College group had formal training programs for supervisors in how to conduct appraisals. It appeared that most supervisors were supposed to learn how to do appraisals by informal methods. On the other hand, four of the U.S. Schools had some form of appraisal training although it was optional in two sites and subject to budgetary cutbacks. The three other U.S. School sites relied on detailed procedural manuals and informal learning.

All the U.S. Schools, and some of the businesses and colleges, had formal introduction and briefing sessions for administrators and supervisors when new performance evaluation programs were introduced. However, the level of maintenance training or training for subsequent new supervisors rapidly dropped off. These findings raise some doubts about the thoroughness of the commitment to effective performance appraisal programs which exists at senior levels of management, or suggests a strong belief in the efficacy of informal learning about a system.

Finally there is also the question of training of subordinates in the appraisal process. Conceivably, subordinates should be equally well trained in how the system is supposed to work and what is to be expected of them in the appraisal process. In general, the three groups of organizations did try to convey to subordinates the sorts of performance expectations that would be used for appraisals and the formal methods by which information on performance would be gathered.

This orientation or training allows subordinates to be somewhat self-managing in that they know to some degree what is expected of them and how these expectations will be evaluated. This is consistent with the notion that professionals should work in a fairly independent manner, largely self-responsible for the guidance of their conduct. What is missing however, is training in how subordinates should behave in the processes of evaluation; i.e., the setting of goals and action plans, the development of performance criteria, the preparation of self-evaluation and the discussion of development plans. How to perform well in these sorts of activities did not seem to be the subject of formal training in any of the study sites and it seems to be something that has to be learned informally through interaction with the supervisor. Thus one more role burden is suggested for supervisors: training and counselling of subordinates in how to go through the appraisal and evaluation process.

3.3.7 Special Cases and Organization Justice

The detailed arrangements for a special review of performance problem situations and for provision of organization justice vary widely from site to site, but the general principles are much the same. The elaboration of procedures is normally much greater in the unionized colleges and schools than in the business organizations. However, the personnel departments in some businesses have developed equally detailed procedures in the absence of a union. In addition, legal trends are forcing all organizations to be much more careful and thorough in dealing with problem situations than in the past. The results in this section will be discussed in two parts: probationary employees and established employees.

a) Probationary employees: Almost all of the organizations in this study provided a fairly stable work environment with job security or quasi- or formal tenure available to persons with satisfactory performance evaluations. For most of the organizations in this study, a psychological or implied contract exists between the employer and subordinates that once the probationary period of employment is satisfied, then job security is essentially assured.

Along with this implied contract of quasi- or formal organization tenure, is also the implied commitment of the organization to using internal sources of human resources for most higher level staffing needs (i.e., for all supervisory and managerial jobs but excluding specialists and perhaps very senior executives.) Thus it is incumbent on the organization to be very careful about who is recruited to work for the organization and, especially, who is allowed to stay after probationary employment.

Six of the U.S. Schools have quite detailed evaluations of probationary teachers. The content of the evaluations focus on both subject area knowledge and competency on whatever criteria for good teachers have been established by the school jurisdiction. Formally or informally, the school jurisdiction creates a certification process in addition to the teaching certificates one might earn from a teachers' college. The probationary period of employment lasts until the completion of the certification process.

Once the person has been assessed as qualified or competent (within the allowed time limits), then the teacher is in effect licensed to practice in that school jurisdiction. The seventh U.S. School site is a school board which accepts teacher certificates issued by the State for admission to probationary employment and then looks to see if a teacher receives satisfactory performance ratings during the probationary period. There are no special criteria, detailed appraisals, or attempts to create local certification for probationary teachers in this seventh site.

Periods of probationary employment may last from one to two years and opportunities for repeat tests or trials may be provided, but the ultimate result is the same. If a person fails to qualify in the mandatory probationary period, then the teacher is not allowed further employment in that school jurisdiction. Opportunities to receive counselling and coaching are generally available for teachers who are having difficulty qualifying, but the conditions of dismissal or acceptance are clear and objective. Thus there are no legalistic disputes concerning dismissals as a result of failing to qualify during the probation period.

The colleges also pay very careful attention to instructor performance during probation because of the concern for instructor quality, and the fact that few instructors have received prior training as teachers. Some of the colleges have established sets of required instructor competencies similar to the U.S. Schools, but there is no attempt to do formal local certification or licensing of instructors. Instead, the probationary instructor has to receive satisfactory course and student evaluations during the probationary period and if successful, is hired as a member of the "permanent staff".

In two of the Business group the minimum standards and criteria for probationary employees were well established and failure to qualify was treated similarly to the schools and colleges. The other three businesses had less definite performance expectations for probationary employees and consequently used more individualistic judgements about whether a person qualified for "permanent" status. Generally, if probationary individuals received satisfactory evaluations in the same appraisal process used for all other professionals, they were eventually considered to have passed from probationary to permanent status. Informally, a one year probationary period was used with opportunities for qualifying possibly extending for another one to two years. In summary, the U.S. Schools are the most formalistic and careful in probationary performance appraisals, the colleges are less formalistic but equally careful, and the businesses less formalistic and less careful.

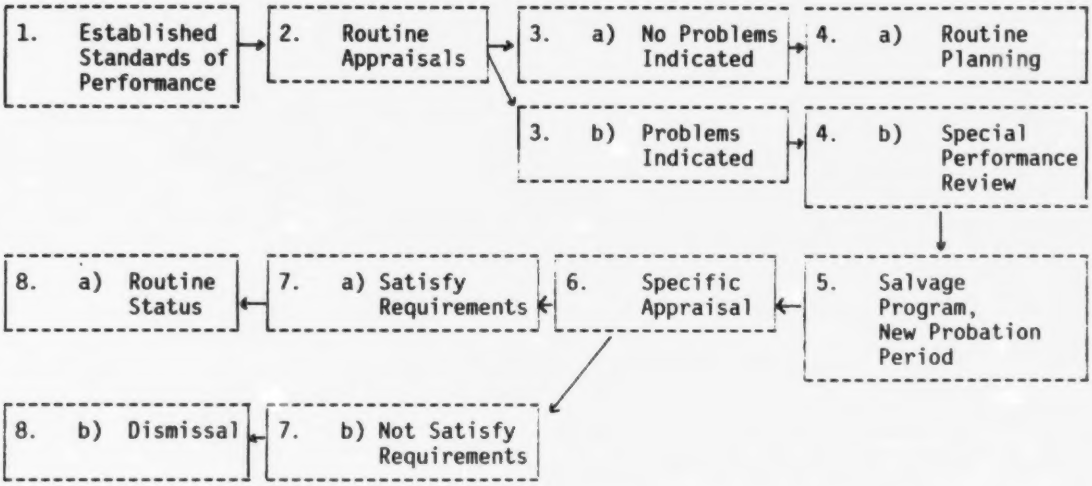
b) Established employees: This term is used to describe employees who have satisfactorily completed probation and are now regarded as essentially permanent employees of the organization. Their tenure may be formal or informal and as long as they continue to receive some satisfactory minimum rating in performance evaluations, they continue to accumulate seniority in the organization. If established employees start to receive less favourable performance evaluations, then a special review process or renewed probationary status is triggered into action. This was the case in all of the organizations in this study. The detailed procedures of the special review process varied from site to site but the essential features of the systems were the same.

The essential features of the special review process are as shown in Figure 1, which outlines a model of a performance-monitoring system. The performance-monitoring system is easier to manage if performance standards have some clearly established minimum levels and criteria. If the minimum acceptable levels of performance are nebulous, then it is much more difficult to identify performance problems. Most of the schools and colleges were more careful to establish minimum required standards and two of the businesses did so as well. Those who left these criteria unspecified were subject to contentious debates when attempts were made to confront performance problems.

Salvage programs could go by many names but they all involved supervisor-subordinate discussions of indicators of unsatisfactory performance, analyses of causes, coaching and/or corrective plans, and subsequent specific appraisals on the results of the plans. Eventually

Figure 1

Example of a Performance Monitoring System with Special Performance Review


1. Established Standards of Performance
2. Routine Appraisals
3. a) No Problems Indicated
4. a) Routine Planning
3. b) Problems Indicated
4. b) Special Performance Review
5. Salvage Program, New Probation Period
6. Specific Appraisal
7. a) Satisfy Requirements
8. a) Routine Status
7. b) Not Satisfy Requirements
8. b) Dismissal

a decision has to be made as to whether inadequate performance is correctable or has been corrected, or whether the employee is judged to be incompetent for the job requirements. If the level of performance persists at an inadequate level despite the best efforts of a salvage program, then the evidence points to incompetence as the cause.

Evidence for demonstrated incompetence (provided by the monitoring and a special review process) has generally been accepted by the courts and arbitration systems as sufficient grounds for dismissal for cause. Other options used for dealing with incompetence were transfer of the individual to a less sensitive or demanding job, redesign of a job so as not to require the deficient competencies, and to have frank, career planning discussions with problem employees about whether they really belong in the particular profession or would be better suited to a less demanding organization. Due process, fair treatment, and rigorous documentation of all steps and procedures in special performance review cases are necessary for the protection of both parties.

All the sites in the study have developed systematic and carefully documented procedures for whenever serious performance problems occur among established employees. In practice, the actual use of such procedures is comparatively rare. One reason is that rigorous screening at the probationary level removes most potential performance problems at the outset. Another reason is that the special performance review process is time-consuming and costly and the organizations have to go through some very careful consideration of whether a performance situation is sufficiently serious to warrant the procedure. Finally, counselling sessions with inadequate performers often result in voluntary withdrawal from the organization. Employees with a performance problem may also not wish to go through the complications and tensions of the special review and salvage program process and, instead, will try to seek alternative employment once they know where they stand in the present organization.

One difference which was found between the Business group and the educational organizations concerns employees who consistently receive performance evaluations at minimum satisfactory levels. Technically these employees satisfy the conditions of employment, but they are marginal or mediocre employees in the eyes of their employers. The schools are basically obligated to leave these teachers alone if they are minimally satisfactory because education organizations have less discretionary latitude for dismissing marginal performers than do business organizations. The colleges in this study try to avoid a situation of marginal performers by requiring a high level of teaching performance as a minimum acceptable standard during the probationary period. If what might be called "teaching excellence" is demonstrated during probation, the colleges hope that long-run performance will continue at an equivalent level.

Most of the Business group have the philosophy of trying to require excellence as a minimum standard of performance. Persons who consistently receive marginal, barely acceptable, or below average evaluations are not assured of job security in these organizations. Professionals with this sort of record receive the lowest possible salary increments, do not receive promotions, and have to participate in counselling discussions about their suitability to work in that organization.

3.3.8 Evaluation and Maintenance of the Appraisal System

The evaluation of an appraisal system is essentially a matter of program evaluation, that is, an assessment of the outputs and consequences of a program in comparison to the goals established for the program. For the organizations in this study, it is safe to say that none are conducting formal program evaluations of their performance appraisal systems. This is partly due to recognition of the complexity of what a formal program evaluation requires, but also due to the limitations of resources within the organizations. Most resources are devoted to implementation and administration of the appraisal systems and very little is left over for other purposes. Consequently, any form of program evaluation which does occur is in the form of impressionistic evidence and indirect indicators.

Among the Business group, three of the appraisal systems were felt to be too recently launched for any sort of program evaluation to be done. These systems are operating on management's faith or conviction that the approach is appropriate for their organizations. No revisions are yet felt to be necessary and no record of grievances has yet been established. Indirect indicators such as paperwork burden or sense of frustration with the systems are also not yet apparent. The other two appraisal systems were also launched under the strong conviction by management of the appropriateness of the system and their experience with their systems (both MBO) has not challenged these convictions. There is no particular record of grievances, no frustrations with paperwork, or other indirect indicators of problems. Consequently, the conclusion from the impressionistic evidence is that the appraisal systems in the Business group are operating satisfactorily and no significant changes are planned.

A similar situation prevails in the College group, although a few differences are apparent compared to the Business group. Since some of the colleges used joint task forces to develop their appraisal systems, these systems were launched with the conviction of both managers and instructors that the systems were appropriate. The operations of the appraisal systems are fairly new in all the colleges but the experience so far has not suggested the need for any significant changes. Grievances and arbitrations have occurred in a few situations. These have been concerned with attempts by colleges to dismiss some established instructors for continued inadequate performance. These are very few in number in comparison to the large number of instructors employed at the five colleges. In terms of frustrations with appraisal systems, there are more complaints about paperwork in the colleges than in the businesses. This is interpreted as being due to the greater care paid to probationary evaluation; the greater use of direct supervisor observations and client reports; and the union membership of the instructors. The recent launching of a centralized MBO planning system at the colleges also contributed to the perception of increased paperwork.

Two of the U.S. Schools had a form of program evaluation in that student learning and student achievement test scores were used as criteria for the validation of teacher competencies. Since the teacher appraisal systems were founded on the assessment of competencies, then student performance could act as an indicator of the effectiveness of the appraisal system as well. This type of program evaluation was most extensively developed in the State of Georgia where student achievement records were used both to validate the required competencies and the effectiveness of the teacher certification program. Otherwise, the U.S. Schools followed much the same impressionistic evidence approach to program evaluation as did the colleges and businesses.

Overall, the impressions received from the sites in the study reveal few formal attempts to evaluate the success of the performance appraisal systems and little evidence of any problems with the systems. If problems should begin to develop, then they will probably be recognized only after indirect evidence mounts to a level considered to be intolerable.

3.3.9 Trends and Outlook for the Appraisal Systems

As might be gathered from section 3.3.8, the Business, College, and U.S. Schools groups are basically satisfied with their current systems and may only introduce minor changes to fine-tune or complete stages of introduction of their systems.

All three types of organizations will continue to require employees to maintain a satisfactory minimum level of performance to retain their seniority, although the definition and criteria for "satisfactory minimum" will vary from one organization to another. Each will conduct routine appraisals using various approaches to verify this level of performance requirement. Each will use some form of goal-setting and action-planning in addition to routine appraisals. In the businesses, goal setting and planning will be mostly concentrated on contribution to organization goals with some aspects of professional development planning. In the colleges, there will almost be a balance between professional development plans, contribution to organization goals, and teaching plans. For teachers in schools, goal setting and planning will be dominated by teaching plans, with some professional development plans, and little planning for contribution to organization goals. This latter pattern is somewhat unfortunate for teachers because it implies a role for them as teaching mechanics rather than professional educators. Instructors in the colleges have a more professional role as educators with a more holistic approach to goal setting (and a more holistic approach to appraisals). On the basis of impressionistic evidence alone, in a few U.S. Schools sites where the teachers were seen as being more involved in plans for the achievement of school goals and where individual plans were seen as making a contribution to organization goals, the teachers seemed to be working more truly as professional educators.

4. CONCLUSION

A preliminary conclusion which emerges from both the literature review and the field study parts of the project is that a distinction is not consistently made between appraisal and evaluation of performance. The two terms are used interchangably in many cases but the following distinction will be used in the conclusion of this report. The term "appraisal" will be used to refer to the formative process of gathering data about performance and giving feedback, provided that the feedback is objective or non-judgemental.

On the other hand, "evaluation" will be used to refer to judgements or decisions about the relative level of performance and the giving of feedback about the judgements. A key aspect of the difference between appraisal and evaluation is that in appraisal, it is the subordinate who has to make judgements about whether the behaviour was effective/ineffective. In evaluation, it is the supervisor who makes these judgements. If both parties share a common understanding of what constitutes effective/ineffective performance, then appraisal will be used more than evaluation and subordinates can be more autonomous in their work.

The appraisal and evaluation of performance of professionals has posed considerable difficulties for many organizations for at least three major reasons. One reason lies in the nature of professional bureaucracies where professionals perform their service work in a largely autonomous manner. As a result of their professional training (and in many cases, licensing or certification), they are assumed to be competent for the autonomous performance of their duties and assumed to be able to act in a self-regulating manner. Under these assumptions, individual accountability for professional performance is only subject to question when extreme exceptions occur to accepted norms of professional conduct. Performance which is an exception to these norms is outside the commonly shared understanding of what constitutes acceptable performance and the delegation of autonomy is revoked.

The second reason for difficulty in appraisal and evaluation of professionals is that the means and techniques (process) of a professional's work and the results (output) may be difficult to measure or assess objectively. If the work of a professional is preventative in nature, then the absence of measurable outputs may be a more valid means of appraising performance. If the results of a professional's work are qualitative or long-term in nature (i.e., to enhance feelings of self-esteem in a client; to create a sense of citizenship, and and so on), the results can be extremely difficult to assess as well. However, one should keep in mind the observation that "if the results of someone's work cannot be measured, rated, assessed or even qualitatively described in some manner to distinguish between when the work is or is not done, then we have to wonder whether any work is being done at all".[22]

The third problem lies with the assumption that a direct correspondence exists between the work of the professional and the results which are eventually obtained. In situations where numerous other variables can intervene to affect end result outcomes, this assumption can be of questionable validity. In face of the uncertainty about a performance-results correspondence, many professional bureaucracies focus appraisals on the process of professional work. If the processes are felt to be appropriate, then it is hoped that the work of the professional will be a positive contributing factor (not necessarily causal) to the desired end results.

Despite these principal difficulties, professional service organizations still have to grapple with the problem of appraisal in order to continuously try to develop or maintain a high level of professional service and to have a valid data base for personnel evaluation. This means that some form of systematic appraisal based on some means of obtaining valid data about performance has to be developed for each organization. Once the information about performance has been gathered, the second issue is how to make evaluations or judgements about the performance. To evaluate performance, it is necessary to have some sort of standards for purposes of comparison. Normally, evaluation requires the definition of some minimum standards of performance for at least the key aspects of a professional's work, so that the work of all persons can be judged in comparison with some fundamental requirements of the organization. Some organizations may wish to go further and define a range of standards above the minimum so that performance may be evaluated with respect to a hierarchy of standards. This will depend upon the organization's philosophy or policies about promotions and careers for professionals or the policy about performance-related compensation differentials.

---

[22] Odiorne, G. "What's Wrong with Appraisal Systems?" In Personnel Policy: Issues and Practices. Columbus, Ohio: Merrill Books, 1965.

4.1 OBJECTIVES AND PURPOSES OF AN APPRAISAL SYSTEM

In order for an organization to make an intelligent selection of an appropriate mix of appraisal methods, it is necessary for the organization to decide on priorities of the purposes of the appraisal system. The recommended general principle is that there should be consistency between the type of appraisal and the purpose being served by the appraisal. For example, if the organization has a very strong belief in the efficacy of a particular approach to teaching, then the appraisal system should be activities-based to ensure that the desired methods are being used. If the organization has a variety of purposes to be served by appraisal then a variety of appropriate methods will have to be used to form a multi-methods package for appraisal and evaluation of performance.

All of the organizations in the field study had some form of multi-methods appraisals. No one approach to appraisals was found to be appropriate for all purposes and each combination of methods had its strengths and limitations. Some of the multi-method appraisals were dominated by competency and activities-based appraisals, others were dominated by various forms of results-based appraisals. The range of approaches was equally varied between businesses, colleges, and schools.

While it seems self-evident that an appraisal system should be developed to serve organizational objectives and purposes, it is less self-evident which type of system serves which type of purposes. A purpose which may be common to all organizations is the need for an appraisal system to provide an objective, valid, and reliable data base for routine personnel administration and decisions. Over and above these purposes, the need is for an appraisal data base to provide information for special purposes or situations in an organization which may be different from other organizations. The following sections represent some of the special purposes which were found in the field study.

4.1.1 Situational Factors

Some of the U.S. Schools were in situations of rapid population growth where the demand for teachers exceeded the normal supply. This meant that several teachers had to be recruited from out-of-state, where little was known about different qualities of teacher training. Secondly, the quality of teacher training within the state was known to vary greatly as well.

In another case, some U.S. Schools were confronted with a crisis of a lack of confidence in the school system's ability to deliver a quality education. Since enrolment was static or even declining, it was not possible to use the approach of trying to change the quality of inputs of new teachers and the school system had to work with available resources.

A third type of situation found more in colleges and businesses was the need to develop or maintain a competitive distinction in quality of professional service. The situation faced by these organizations was that if they slipped or deteriorated in professional excellence, their clients/students could easily go to competitors who provided better service/education. A fourth situation for a number of organizations was slow or static growth with reasonably high standards already achieved. For organizations in these situations, the challenge was to both maintain the level of quality of service and provide for continued professional development of staff.

For these different organization situations, serious consideration has to be given to the question(s) of what needs or purposes are to be served by the appraisal system. For schools in high growth situations, the focus of appraisals was on controlling the standards of entry of professionals into the organization. For schools in near-crisis situations, the focus was on specific re-direction of available resources to different methods and priorities in educational services. For organizations in highly competitive situations, the focus was on excellence of service and market sensitivity for competitive advantage. For stable, good quality schools, the focus was on maintenance of standards and continual upgrading and development. These patterns are not meant to suggest that only these combinations of situations and approaches to appraisal will be suitable. Instead, the conclusion is that the focus of the appraisal system should be relevant to the situation faced by a particular school jurisdiction and that the appraisal systems may need to be reorganized from time to time to maintain a match with evolving organization situations.

4.1.2 Sense of Mission and Normative Models of Professional Service

An explicit (or strongly implicit) sense of mission in an organization or normative model of professional practice has the effect of defining a number of priorities for resource allocation and boundaries for acceptable and nonacceptable activities and behaviours. In effect, the organization creates a particular culture which then influences behaviour within the organization community in a manner analogous to the influence of a societal culture on a village. If the organization has a strong conviction that a particular method of professional service will provide the best results,[23] satisfy a purpose, or achieve a mission, then the appraisal system can be designed to elicit behaviours which fit the normative model. Normally this is done by use of an activities-based or process appraisal system, but can also be done using a results- or quasi-results-based appraisal system. The key question about a normative model is whether it has any validity for actually being an effective or more effective method of performing some professional service. Conceivably, an appraisal system could carefully monitor professional activities for correspondence to the normative model, but the organization goals or mission may still not be achieved if the model is unsound.

---

[23] See for example, Skinner, B.F. "The Shame of American Education". American Psychologist 39, no. 9 (Sept. 1984), pp. 947-956, which advocates programmed learning.

Among the organizations in this study, only a few had explicit normative models for professional services which were based on theories published by respected authors in the field. Almost none attempted any sort of validation research to see if professional service corresponding to the theory actually generated better professional service or results. Most organizations which used some sort of normative models based them on face validity, i.e., "common-sense" or collective judgements by task-forces about what constituted or was necessary for good professional service. If these collective judgements are valid, then appraisals based on these models can also be valid.

The conclusion here is that to the extent that an organization has an explicit mission or normative model for professional service, the appraisal system can be used as a powerful technique for achieving the mission or implementing the normative model. However, it is necessary for an organization to first clarify whether there is a strong sense of particular mission for which collective energies should be marshalled, or a strong conviction about a normative model of professional service to which collective energies should conform. If either of these situations is true, then the appraisal system could be designed to specifically contribute toward the achievement of these purposes.

4.1.3 Career Stages of Professionals

Most of the organizations in this study had an implicit system of different approaches to appraisal for different stages in a professional's career while an explicit distinction existed between types of appraisal for professional service and supervisory levels of the hierarchy.

The conclusion here is that what works informally and effectively in appraisals in many of the field site organizations could also work more formally as a means to reduce paperwork and increase the relevance of appraisals. To this end, an organization could use a comprehensive and multi-faceted appraisal system in recognition that some aspects will or will not be relevant to different persons, or change in relevance for individuals over time. Thus if a supervisor established that a person has satisfied performance requirements for one level of professional service work, the person is assumed thereafter to be competent in this dimension of performance unless evidence turns up to suggest the contrary.

A comprehensive, multi-faceted, and selective approach to appraisals can capture the wholeness of a professional's work and yet avoid redundancy of inappropriate appraisals. The system works best for those individuals who are satisfactory or better performers and who continue to make progress in the organization. If people fail to maintain standards or make required progress, then the more comprehensive parts of the appraisal system come into play again and, in effect, a special performance appraisal system is established for that person.

4.1.4 Absence of Mission or Mormative Model

Despite the importance of having an *organization mission* and normative model for professional service, many organizations will not be able to specify a normative model. Instead, the organization may propose a very general mission, such as "provide good quality education" and remain non-specific on what is good quality education, or how it is to be achieved.

Two strategies for appraisal seem to be commonly used in such situations. The first strategy is to try to recruit and select the best possible candidates at the outset and turn them loose to do the best that they can. If the majority of the human resource inputs into the organization are the best possible quality, then it is assumed that they will somehow generate good results from their collective effort.

The second strategy used in the absence of explicit mission or normative models is to delegate these decisions to the professionals themselves. The strategy of delegation of decisions on missions and methods to professionals can work if the organization has a high degree of confidence in the professional component. If either or both the trustees who provide the financial resources or the clients who receive the professional service decide in some way that they are satisfied with the services, then confidence is established for delegation of self-regulating responsibility to the professionals and administrators.

If a crisis of confidence occurs, then the community, through a variety of representatives, must act to reassert responsibility and authority. This relationship of community vis-à-vis a profession must hold for virtually any profession because the community provides the resources for the profession and the ultimate justification for any profession is service to the community.

If definition of normative models of professional service is delegated to the professionals themselves, appraisals can be greatly clarified if the professionals achieve a consensus on the normative model. However, a common situation in professional bureaucracies is that little consensus exists among professionals about the most effective methods and approaches to provide professional service. Consequently, it is difficult for the professionals to be self-regulating if they have significantly divergent views about what constitutes good professional service or results. Appraisal in these circumstances reverts back to impressionistic judgements about "good" performance based on non-specific, subjective, and non-aggregated criteria.

4.2 DESIGN, IMPLEMENTATION, AND SUPPORT OF APPRAISAL SYSTEMS

Once questions have been answered about purposes, missions, and so on, of appraisal systems, the next stage, chronologically, is design and implementation of the system.

Generally there were two approaches to design and implementation used by the organizations in this study, the top-down approach and the consultative approach. In the top-down approach, the purposes, design, and implementation of a system are determined by senior management with the aid of the personnel department and conveyed to employees as part of the conditions of employment in the organization. This approach was used by most of the businesses and about half of the colleges and schools. In the consultative approach, senior management invites the participation of subordinates in various forms of joint-committees to work on the development of a system. Usually, in the consultative approach, senior management has decided on the major purposes of a system, but invites consultation on the design and implementation, particularly for the operational details. The consultative approach was used by a few businesses and the rest of the colleges and schools.

In the consultative approach, senior management made the fundamental decisions on either basic purposes, missions, or long range goals to be served by an appraisal system, while the decision-making about operational aspects was delegated to local school boards, departments, branches, and so on, for implementation. In other words, the central authority defines a common core of appraisal aspects which have to be satisfied by all systems, but leaves definition of the other aspects up to local joint-committees for decision making. If the common core is very comprehensive and specified in great operational detail, then a top-down system is essentially being used because very little scope is left available for local decision-making. If the common core is essentially a statement of principles and philosophies with little operational definition, then a great deal of scope is delegated to local decision-making. The scope and comprehensiveness of the common core is a matter determined by the degree to which the centrality authority is willing to trust or has confidence in the local decision makers.

In terms of implementation, all of the organizations developed extensive manuals and procedures for orientation of supervisors and subordinates to the appraisal system, as well as guidance in conducting appraisals and documenting the process and the outcomes. However, most training was concerned with familiarization of supervisors and subordinates with the formal manuals and procedures so that implementation concentrated on technical knowledge. In a few cases, procedural training was supplemented by role-playing exercises where supervisors could practice the application of their technical knowledge. Training tended to be for the introductory launch only or for newly promoted supervisors. There was very little evidence of follow-up training, except on a voluntary basis in some organizations. Maintenance of the systems seemed to depend in many cases upon the personnel department's administrative procedures or demands for periodic appraisals and pressures for documentation. In some organizations, the outcome of this form of maintenance was the gradual development of the attitude that performance appraisal was done to satisfy the personnel department or some centralized bureaucratic procedures.

The conclusion is that implementation of an appraisal system was more effective when the system was seen as being either locally or jointly developed (local ownership), or a management system (management owned) rather than a personnel department program (bureaucratically owned).

4.3 METHODS AND APPROACHES TO APPRAISAL

The conclusion about approaches is that there is no one best approach and that most organizations will have to use some form of multi-faceted approach. Which combination and which emphases will be chosen will depend upon the needs, priorities, and purposes of each organization.

The disadvantage of a multi-faceted appraisal system is that principals will have to spend more time thinking about and discussing what types of appraisal are relevant for individual subordinates, and this will be a burden to principals with scarce time resources and wide spans of control. Many principals, for good reason, prefer the simplicity of a standardized checklist form of appraisal.

The advantage of a multi-faceted system is that when principals do spend more time thinking about and discussing an individual's performance and deciding what is relevant in appraisals, then the quality of individual supervision and management of human resources in a school is likely to be enhanced. Principals would have to think about different teachers as individuals, with different strengths, weaknesses, needs, and stages of development. If serious dialogue occurs between principals and teachers on these issues and if appraisals are tailored appropriately, then a considerable improvement in the quality of supervision and a more effective utilization of human resources should result. Furthermore, the problem of redundant or non-relevant appraisals should also be avoided for areas of performance where satisfactory standards have been well established.

4.4 DIMENSIONS AND CRITERIA FOR APPRAISALS

As previously stated, the foundations for the use of any set of performance dimensions for appraisal should be some theory, philosophy, or normative model of good quality professional service. Once some sort of normative model has been developed, then related performance dimensions and required competencies can be identified. The competencies should identify who can do the professional work, and the performance dimensions should identify the types of behaviours that produce desirable professional work.

The key aspect to performance dimensions and criteria is that they have to be valid representations of the performance the organization had defined as being desirable or had chosen as a normative model. If an organization truly believes in the efficacy of its particular normative model, then it should be possible to also specify some results criteria which should be observable or testable from the use of the model. If the results criteria are actually produced, then the use of the model is justified or the validity of the model is established. In practice, very few of the organizations in the field sites had a well-developed or explicit normative model of professional behaviour as a foundation for their appraisal systems, and even fewer used results criteria to verify their model.

The second challenge to development of criteria is to describe or find indicators of a range of performance for each dimension of performance. A particular criterion may consist of a two-point scale (anything below is unacceptable, anything above is acceptable) or some form of continuum (usable for development purposes or to distinguish relative levels of performance). If the criteria are highly explicit so that performance can be rated accurately and with reliability by external observers, then the criteria are said to be objective. Where criteria are not explicit for a performance dimension or cannot be rated with reliability, then the criteria are said to be subjective. While everyone seeks to develop objective criteria for performance, the reality is that many criteria are subjective and hence subject to dispute. Given this reality, the key issue seems to be acceptability of criteria more than objectivity. Where there is low acceptability about a performance dimension or how it is measured, then there is a strong need felt for objective criteria for performance. Where there is mutual confidence and acceptance between supervisors and subordinates, then subjectivity in appraisal criteria is more acceptable.

The conclusion here is that performance dimensions and various forms of criteria can only be compromises of varying degree. No list of performance dimensions can be comprehensive, and the best that can be hoped for is that the list will be reasonably representative of the important dimensions. Secondly, few criteria will satisfactorily represent the range and complexity of performance, especially for professional service types of work. Consequently, the realistic limitations of an appraisal system, no matter how complex and detailed, have to be recognized. Perhaps the most that can be expected or hoped for is a system which manages with accuracy at a level better than chance, to distinguish between acceptable (satisfies minimum standards) and less-than-acceptable performance. In addition some organizations may also wish to distinguish between performance which is acceptable and performance which is above the average.

4.5 SOURCES OF PERFORMANCE DATA FOR APPRAISALS

The review of the literature and the field study results considered four basic sources of information about the work of professionals: self-reports by professionals, client reports and results, peer reports, and supervisor monitoring and observations. All of the sources of information are valid in the sense that the work of a professional impacts on all of these components of the organization and each component is a stake-holder (i.e., affected by) in the work of the professional. However, the frequency of use of the different sources of information was quite mixed in actual practice.

Self-reports seem to be normally used for individuals with experience or seniority who are relatively well establshed as effective performers. The use of self-reporting is a mark of recognition for the established competence or capability of the person to be largely self-managing as a professional. Conversely, some organizations used short-term self-reporting for individuals in special performance or renewed probation situations to force the problem performer to confront the need to take some responsibility for his or her performance.

Reports and results by clients/students and subordinates were commonly used in very market-oriented businesses and colleges as an important source of information for appraisal. The use of these sources is consistent with the mission of these organizations to be market-oriented and to succeed by providing excellence of professional service.

Supervisor monitoring and appraisals, using both formal and informal information, was the most common source of data for appraisals in all of the organizations. The key role of the supervisor in the appraisal process has been strongly pointed out in a number of parts of this report. It seems to be inescapable that the ultimate appraisal of a professional's performance rests with the information gathered and judgements made by supervisors. This conclusion holds as much for an appraisal system that is highly formalistic, legalistic, and "objective" as for a system which is informal, subjective, and non-specific. Consequently it follows, and cannot be stressed too strongly, that supervisor training at all ranks in the organization is crucial to the successful functioning of an appraisal system.

Each organization has to choose which sources of information are to be used in addition to supervisor monitoring for the gathering of data for performance appraisal. The use of peer appraisals are appropriate for development purposes and where the regulation of professional services is delegated to the professionals themselves. Self-reports are appropriate for development purposes and where activities-based and results-based appraisals are used. Client reports/results are appropriate where a strong sense of mission/normative model exists or where program evaluation is desired to verify the results of a mission or model. None of these approaches or purposes are mutually exclusive and a suitable mixture will have to be chosen to satisfy the various needs and priorities of the organization.

4.6 STANDARDS FOR EVALUATION OF PERFORMANCE

As previously discussed, a distinction is made between appraisal (gathering information and providing feedback) of performance and evaluation, which concerns decision-making about the acceptability of performance. At the most basic level, evaluation is concerned with the decision about whether performance satisfies minimum acceptable standards. Normally, this type of evaluation occurs for two phases of a person's career; minimum standards to satisfy probation requirements, and minimum standards to satisfy requirements for continuity of employment or maintenance of seniority. Minimum standards are not necessarily low standards and can evolve over time as environmental conditions and organization goals change. A second type of evaluation is concerned with distinguishing between levels of acceptable performance or to distinguish between above-average and average performers in the organization.

From the results of this study, it was clear that evaluations depended upon the ability of supervisors to make judgements about performance which reduced complex appraisal information into simplistic summaries about overall performance levels. To the extent that an organization uses such evaluations for personnel decisions (compensation, training, promotions, renewed probation), this places another heavy burden of responsibility on supervisors to make complex judgements. The literature review and some of the field site results show that many supervisors try to cope with this responsibility by evaluating most subordinates as acceptable or average. This practice has turned out to have unfortunate subsequent consequences for some organizations. The conclusion once again is that supervisor training and organizational support is crucial for effective and fair evaluations of performance.

Another caution about evaluations of performance is that they should be appropriate for the needs of the organization for management of human resources. This was a significant difference found between the business and the educational organizations. It was very important in the Business group to distinguish between levels of performance of individuals because these evaluations were the determining factor for personnel decisions about performance-based compensation, job assignments, and promotions. These sorts of human resource management decisions were far less common in the colleges and schools, although more prevalent in the colleges. Consequently, there was far less need in the educational organizations for performance evaluations to distinguish between levels of performance or to try to identify potential for future types of performance.

Minimum standards of performance for administrators may be very difficult to establish because of the variety of administrative responsibilities and the fact that so much of an administrator's efforts are devoted towards negotiated goals. Since these can change from year to year, there is less stability of content to an administrator's job than a teacher's job. One way to approach the question of evaluation for a supervisor's job is to think of it in terms of three major components:

a) routines which must be satisfactorily done year in and year out,
b) problems which are exceptions to routines or occur outside of routines and which must be dealt with as they arise,
c) changes which are to be worked on as parts of longer range goals and planning activities.

A reasonable minimum standard of performance or the minimum level of accountability to be expected of administrators should be the accomplishment of routines for which that job is responsible. Thus no matter who is the incumbent in the position, the organization has the right to expect that the routines will be satisfied.

Some individuals will have a better record at dealing with problems and exceptions than others, but the organization has the right as well to expect administrators to have some degree of competence in this aspect of management. What sort of evaluation is made of a person's performance in this area would depend upon the effort by the individual and the degree of seriousness of the problem or exception.

Finally, performance in the area of changes or goals represents either specific commitments by administrators or specific assignments to administrators. These aspects of performance represent responsibilities or accountabilities which are different from routines and which act to individualize a person's job description. Since goals and plans are specific commitments by individuals, it is appropriate for them to be held accountable for the results (or quasi-results). The evaluation of performance in this area will depend upon the assessments made about degrees of progress towards various goals and relative importance of different goals.

4.7 SPECIAL REVIEW AND RENEWED PROBATION

The results for special performance review and renewed probation are relatively consistent for all the organizations studied. While the details and formats may differ from site to site, the general principles are the same. The personnel policies at each organization endeavour to provide for due process and fair treatment of each individual found to have a performance problem. Salvage programs are used where individuals receive feedback and counselling and are provided with opportunities to recover their former satisfactory level of performance (or build up to a required performance level if new standards are established). Progress during the period of renewed probation is carefully monitored and time limits are established to allow a fair but finite opportunity to recover. If the individual apparently cannot satisfy the minimum required standards despite repeated trials and salvage programs, then the individual has demonstrated lack of competency to perform the required type or level of professional service.

Differences between the study sites were in three areas of practice. Some organizations had established more explicit minimum standards of performance than others with the effect that individuals could be more self-monitoring and recognize when their performance was near a danger zone. Some organizations allowed longer time periods for renewed probation or salvage programs with the effect that the individual had a longer time available to recover or re-establish necessary performance levels (or a longer time to avoid doing anything). The organizations which tended to be more explicit and demanding in performance standards and quicker to take action on performance problems, were those who tended to be more market-sensitive or with firm policies of professional service excellence.

The conclusion here is that if the organization has a firm commitment to a strategy of market-sensitivity or a firm commitment to excellence of its professional services, then a number of policies and practices in supervision, operational procedures, and personnel policies have to be consistent with the goals or act to give force to these goals. The approaches taken by an organization in requiring high standards of performance, being explicit about standards, and taking care to maintain standards are an example of this need for consistency. An organization which operates in a quasi-monopoly situation or enjoys secure access to resources might not feel such a need for market sensitivity or accountability for use of resources and, hence, could afford to be less explicit and demanding about levels of professional service offered to the public.

4.8 PROGRAM EVALUATION AND MAINTENANCE

In section 4.1 it was stressed that an organization should try to define goals and purposes to be served by a performance appraisal program. Secondly, it was stressed that goals and purposes should be related to the special needs or situation faced by an organization in relation to its environment. The subsequent topics of design, methods, and implementation were all concerned with the creation of an organization mechanism (the appraisal system) for the achievement of, or the direction of effort towards, those goals and purposes. Program evaluation is concerned with completing the cycle by examining whether the appraisal mechanism does in fact work towards or achieve those goals.

While it would seem that program evaluation is a crucial activity in the management of an appraisal system, and is consistently recommended in the literature, it was actually found rarely in practice. The majority of the appraisal programs were informally evaluated by managers' subjective and impressionistic evidence about whether a program was "working well" or relatively free of problems. If programs were felt to be satisfactory as a result of such impressions, then it was assumed that progress was being made toward the achievement of organization goals.

This can be a risky approach to evaluation of a program because satisfactory progress on established procedures and lack of dis-confirming short-term evidence are only suggestive of or consistent with achievement of goals and purposes, but cannot be equated with such achievement. Consequently it is of great importance to follow through with program evaluation to ensure that the organization's resources are not being wasted or that the appraisal system is actually fulfilling its role. The fact that so few of the organizations were using a formal approach to program evaluation suggests that either they do not realize the risk of such an omission or that they have great confidence in the appropriateness of their particular approach to appraisals.

4.9 SUMMARY CONCLUSIONS

In summary, three general conclusions are derived from the results of this study.

a) Different approaches and priorities in appraisal systems are used for relevance to the local situation of a particular organization or school jurisdiction. This requires all participants in the development of a system to carefully consider the needs and characteristics of their situation and carefully develop goals and purposes appropriate for their situation.

b) Appraisal systems can be expected to go through stages of evolution as all participants have to go through a learning curve of experience with a system. Therefore no one design can be expected to be correct at the outset nor remain in force indefinitely. Different approaches to appraisal may have to be used for different stages of individual career cycles, different levels of jobs in the hierarchy, and for different levels of experience with a system.

c) An appraisal system developed jointly between supervisors and subordinates has a better chance of incorporating diverse but relevant points of view than a system developed centrally by top management alone. The greater the opportunity for participation by parties affected by a decision, the greater the potential quality and acceptability of the decisions. However, a need also exists in large organizations to ensure that equity, fairness, and minimum standards are provided for across all systems of appraisal. A workable solution seems to be a centrally developed (with joint participation) system of universal standards which have to be satisfied by any appraisal system, and then delegation to local authorities (or joint committees) for the development of local, operational specifics.

In the end, no system of appraisals can be workable, no matter how carefully designed, if the people who have to work with it and live by it do not view it as being acceptable and legitimate. Representative participation in the development of a system was not found to be absolutely necessary for acceptability and legitimacy of a system, but participation was found to greatly enhance acceptability. The other side of the coin is that the participants in the joint committee have to share in the goals of an organization or share in the commitment to professional service. If various participants in the planning process use participation as an opportunity for the pursuit of other political goals, then the necessary mutuality of goals will not be present and decision quality and acceptability will not be enhanced. This caveat aside, the probability of developing an effective and appropriate appraisal system is greatly improved through the use of joint planning and development of the system.

APPENDIX 1

CONDUCTING APPRAISAL INTERVIEWS

Step 1: Prepare

- The employee should be contacted several days in advance to allow time to prepare.
- It should be a mutually convenient time.
- Assemble all data (documentation, appraisal instruments, and so on) and refamiliarize yourself with them.
- Ensure that there will be privacy, lack of interruption, sufficient time.

Step 2: Outline the Purpose and Set the Tone

- Make subordinate comfortable.
- Get to the point of the interview (e.g., "As we discussed when we arranged this meeting, our purpose today is to discuss your performance since your last review and what we might do together to help you with any problems you may be having").
- Describe how you intend to conduct the meeting.
- Emphasize two-way communication.

Step 3: Invite Employee's Assessment of His/Her Performance

- Start the discussion by asking "How do you feel things are going on the job? What's going well and what problems are you having? How do you see your performance?" and so on
- If the employee had performance objectives, have your subordinate restate each objective and results obtained.
- Ask questions for purposes of clarification and elaboration.

Step 4: Listen and Learn

- Do not interrupt.
- Listen to what the employee is saying.
- Be prepared to alter your assessment in light of new data.

Step 5: Share Your Data

- Give specific examples of proficiencies and deficiencies.
- Emphasize employee strengths.
- Discuss performance, not personality.
- Tell employees what is wrong with their work, not themselves.
- Avoid comparing the employee to others.
- Allow the employee the opportunity to ask questions, clarify information.
- Explain why you assessed performance as you did.
- When discussing weak areas, focus on solving the problem (causes and possible remedies) rather than on the event.

Step 6: Summarize Your Joint Analysis

- Discuss various aspects of your data and the data of the employee; encourage mutual feedback.
- Praise the positive aspects of performance.
- Review what has been stated, and on what points you agree.
- Discuss where you differ, why you differ (remember complete agreement is not always mandatory), and how to resolve differences relating to future job performance.

Step 7: Develop a Plan of Action

- Emphasize future improvement in performance.
- Encourage the employee to think out solutions to problems (e.g., What do you feel is the problem here? What do you think can be done? How? How can I help?).
- Suggest specific techniques to help improve performance.
- Set formal or informal performance targets and goals (process- or output-oriented).
- Outline jointly what changes are to be expected in the employee's performance in specific terms.
- Review action plans and set follow-up dates to review progress.
- End interview on a positive note of encouragement.

Step 8: Follow-Up

- Summarize the meeting in writing.
- Ensure employee has a copy of appraisal.
- Meet with subordinate on agreed upon date to review progress.
- Make yourself available for coaching and counselling.
- Monitor employee for signs of improvement and give positive feedback.

APPENDIX 2

QUESTIONS ABOUT THE RELATIONSHIP BETWEEN ADMINISTRATOR AND TEACHER APPRAISAL

(1) What proportion of the job functions of teachers and administrators are interactive rather than mutually exclusive?

(2) Should teachers appraise principals?

(3) Should effective teacher appraisal be an objective of administrator appraisal?

(4) How is educational leadership tied to performance evaluation?

(5) To what extent should administrators be evaluated on their ability to evaluate teachers?

(6) Are organizational goals linked to performance at all levels of the school hierarchy?

(7) Should the philosophy and overall objectives of appraisal be the same for both teachers and administrators?

(8) Is it feasible to use the same fundamental process (e.g., number of appraisers, assessment period, data sources, instrument design, goal-setting processes, competency-based measurement) for all school-based personnel?

(9) Do differences in tasks, responsibilities, and criteria for assessment between organizational levels impede the design of a congruent system of appraisal?

(10) Would implementation of the same appraisal process at all levels within the school streamline evaluation? facilitate communication? reduce resistance to the process? serve as a co-ordinating device among levels?

REFERENCES

Acheson, K.A., and Gall, M.D. Techniques in the Clinical Supervision of Teachers. New York: Longman, 1980.

Alexander, E.R., and Wilkins, R. D. "Performance Rating Validity: The Relationship of Objective and Subjective Measures of Performance". Group and Organization Studies 7, no. 4 (December 1982), pp. 485-96.

Appel, J. "Yearly Evaluation - A Key in Assessing Superintendents". Thrust For Educational Leadership (January 1980), p. 28.

Bauman, K.E., and Dent C.W. "Influence of an Objective Measure on Self-Reports of Behavior". Journal of Applied Psychology 67, no. 5 (October 1982), pp. 623-628.

Bazerman, M.H.; Rafik, I.B.; and Schoorman, F.D. "Performance Evaluation in a Dynamic Context: A Laboratory Study of the Impact of a Prior Commitment to the Ratee". Journal of Applied Psychology 67 (December 1982), pp. 873-876.

Bernardin, H.J., and Cardy, R.L. "Appraisal Accuracy: The Ability and Motivation to Remember the Past". Public Personnel Management 11, no. 4 (Winter 1982), pp. 352-357.

Bernardin, H.J., and Beatty, R.W. Performance Appraisal: Assessing Human Behavior at Work. Boston: Kent, 1984.

Blai, B. Jr. "An Appraisal System That Yields Results". Supervisory Management 28 (November 1983), pp. 39-42.

Brush, D.H., and Schoenfeldt, L.F. "Performance Appraisal for the '80s". Personnel Administrator 27 (December 1982), pp. 76-83.

Buser, R.L., and Banks, F.A. Jr. "The Why, What, How and By Whom of Evaluating Principals". NASSP 68 (January 1984), pp. 1-4.

"Can the Schools Be Saved?" Newsweek, May 9, 1983.

Carroll, S.J. Jr., and Schneier, C.E. Performance Appraisal and Review Systems: The Identification, Measurement, and Development of Performance in Organizations. Glenview, Ill: Scott Foresman, 1982.

Caruth, D.; Middlebrook, B.; and Rachel, F. "Performance Appraisals: Much More Than a Once-a-Year Task". Supervisory Management 27, no. 9 (September 1982), pp. 28-36.

Cascio, W.F. "Scientific, Legal, and Operational Imperatives of Workable Performance Appraisal Systems". Public Personnel Management 11, no. 4 (Winter 1982), pp. 367-375.

Cederblom, D. "The Performance Appraisal Interview: A Review, Implications and Suggestions". Academy of Management Review 7, no. 2 (April 1982), pp. 219-227.

Cohen, S.L., and Jaffee, C.L. "Managing Human Performance for Productivity". Training and Development Joural 36 (December 1982), pp. 94-95.

Cordes, C. "Research Finds Little Merit in Merit Pay". Monitor 14, no. 9 (September 1983), p. 1.

Darling-Hammond, L.; Wise, A.E.; and Pease, S.R. "Teacher Evaluation in the Organizational Context: A Review of the Literature". Review of Educational Research 53, no. 3 (Fall 1983), pp. 285-328.

De Bevoise, W. "Synthesis of Research on the Principal as Instructional Leader". Educational Leadership 41, no. 5 (February 1984), pp. 15-20.

Deal, T.E.; Dornbusch, S.M.; and Crawford, R.A. "Villians as Victims: Evaluating Principals". Phi Delta Kappan (December 1979), pp. 273-74.

DeMarco, J.J., and Nigro, L.G. "Implementing Performance Appraisal Reform in the United States Civil Service". Public Administration 61 (Spring 1983), pp. 45-57.

DeNisi, A.S.; Randolph, W.A.; and Blencoe, A.G. "Potential Problems With Peer Ratings". Academy of Management Journal 26 (September 1983), pp. 457-464.

Dittloff, R. "Evaluate the Superintendent". American School Board Journal 169, no. 11 (November 1982), p. 11.

Duhamel, R.; Cyze, M.; Lamacraft, G.; and Rutherford, C. "The Evaluation of Principals" Education Canada 21, no. 2 (Summer 1981), pp. 20-27.

Edwards, M.R. "Productivity Improvement Through Innovations in Performance Appraisal". Public Personnel Management 12 (Spring 1983), pp. 13-24.

Edwards, M.R., and Sproull, J.R. "Rating the Raters Improves Performance Appraisals". Personnel Administrator 28 (August 1983), pp. 77-82.

Farquhar, R.H. "How the Teaching Profession Measures Teaching Effectiveness". Paper presented at the conference on "Teaching in Medical School," Saskatoon, Saskatchewan, June 1978. (ED 1985 683).

Feldvebel, A.M. "Teacher Evaluation: Ingredients of a Credible Model". Clearing House (May 1980), pp. 415-420.

Fisher, C.D., and Thomas, J. "The Other Face of Performance Appraisal". Human Resource Management (Spring 1982), pp. 24-26.

Fleenor, C.P., and Scontrino, M.P: Performance Appraisal: A Manager's Guide. Dubuque, Iowa: Kendall/Hunt, 1982.

Flynn, J.J. "Evaluation of School Administrators". Comment on Education (October 1978), pp. 5-8.

Fordham, A. "Evaluation Reflection: The Context of Teaching and Learning in Studies of Teacher Effectiveness". Studies in Educational Evaluation 8, no. 2 (1982), pp. 111-128.

Gallegos, P.M. "Communicating Performance Results". Journal of Systems Management 34, no. 3 (March 1983), pp. 25-31.

Gelberd, L.B.; Lee, C.; and Lupton, D.E. "Self-Appraisals: The Next Stage in Performance Evaluations". Supervisory Management 28 (May 1983), pp. 9-14.

Glueck, W.F. Personnel: A Diagnostic Approach. Plant, Texas: Business Publications, Inc. 1982.

Gomez-Mejia, L.R, and Page, R.C. "Integrating Employee Development and Performance Appraisal". Training and Development Journal 37 (June 1983), pp. 138-145.

Graves, J.P. "Let's Put Apraisal Back in Performance Appraisal: Part 1". Personnel Journal 61 (November 1982), pp. 844-849.

Graves, J.P. "Let's Put Appraisal Back in Performance Appraisal: Part 2". Personnel Journal 61 (December 1982), pp. 918-923.

Guba, E.G., and Lincoln, Y.S. Effective Evaluation. San Fransisco: Jossey-Bass, Inc, 1981.

Gutridge, B.M. Teacher Competency: Problems and Solutions. American Association of School Administrators, 1980.

Hathaway, W.E. "Testing Teachers". Educational Leadership 33 (1980), pp. 210-215.

Hautaluoma, J.E.; Kaman, V.S.; and Kryger, B. "Performance Appraisal For Human Resource Management: Implications of a Pre-Post Study". Human Resource Planning 6, no. 1 (1983), pp. 41-48.

Herman, J.J. "An Up-Front Approach to Administrative Evaluation". NASSP Bulletin. (November 1978), pp. 34-42.

Hewitt, P.J. "Performance Evaluation: How To Make It Easier on You - and Your Employee". Supervision 45 (June 1983), pp. 14-15.

Holt, M. Evaluating the Evaluators. London: Hodder and Stoughton, 1981.

Hom, P.W.; DeNisi, A.S.; Kinicki, A.J.; and Bannister, B.D. "Effectiveness of Performance Feedback From Behaviourally Anchored Rating Scales". Journal of Applied Psychology 67, no. 5 (October 1982), pp. 568-576.

Huber, V.L. "An Analysis of Performance Appraisal Practices in the Public Sector: A Review and Recommendations". Public Personnel Management 12 (Fall 1983), pp. 258-267.

Hyde, A.C., and Smith, M.A. "Performance Appraisal and Training". Public Personnel Management 11, no. 4 (Winter 1982), pp. 358-366.

Ivancevich, J.M. "Subordinates' Reactions to Performance Appraisal Interviews: A Test of Feedback and Goal-Setting Techniques". Journal of Applied Psychology 65, no. 5 (October 1982), pp. 581-587.

Iwanicki, E.F. "Development and Validation of the Teacher Evaluation Needs Identification Survey". Educational and Psychological Measurement 42, no. 1 (Spring 1982), pp. 265-74.

Kaye, B., and Krantz, S. "Performance Appraisal: A Win/Win Approach". Training and Development Journal 37, no. 3 (March 1983), pp. 32-34.

______. "Preparing Employees: The Missing Link in Performance Appraisal Training". Personnel 59, no. 3 (May/June 1982), pp. 23-29.

Kirkpatrick, D.L. How To Improve Performance Through Appraisal and Coaching. New York: AMACOM, 1982.

Landy, F.; Zedeck, S.; and Cleveland, J. Performance Measurement and Theory. Hillsdale, New Jersey: Lawrence Erlbaum, 1983.

Lewis, A.C. Evaluating Educational Personnel. American Association of School Administrators, 1982.

Lockwood, M., and Ludmer, A. "The Performance Model: An Objective Approach to Measuring White Collar Productivity". Manage 34, no. 4 (October 1982), pp. 30-32.

Lotto, L.S. "Revisiting the Role of Organizational Effectiveness in Education Evaluation". Educational Evaluation and Policy Analysis 5, no. 3 (Fall 1983), pp. 367-378.

Manaktala, V.K. "Job Descriptions: Making Subordinates Accountable". Supervisory Management 28, no. 1 (January 1983), pp. 27-29.

McCleary, L. "Evaluation of Principals". Theory Into Practice (February 1979), pp. 45-49.

McGreal, T.L. "Developing a Teacher Evaluation System: Commonalities of Those Systems that Function Most Effectively". Paper presented at the Annual Meeting of the Association For Supervision and Curriculum Development, March 1982. (ED 226 418).

McGreal, T.L. "Effective Teacher Evaluation Systems". Educational Leadership 39, no. 4 (January 1982), pp. 303-305.

Meyer, J.W., and Rowan, B. "The Structure of Educational Organizations". In Meyer, M.M., ed. Environments and Organizations. San Francisco: Jossey-Bass, 1978.

Mintzberg, H. The Structuring of Organizations. New Jersey: Englewood Cliffs, Prentice-Hall, 1979.

Morrisey, G.L. Performance Appraisals in the Public Sector: Key to Effective Supervision. Reading, Mass: Addison-Wesley, 1983.

Mount, M.K. "Comparisons of Managerial and Employee Satisfaction With a Performance Appraisal System". Personnel Psychology 36 (Spring 1983), pp. 99-110.

Murphy, J.; Hallinger, P.; and Mitman, A. "Problems with Research on Educational Leadership: Issues To Be Addressed". Educational Evaluation and Policy Analysis 5, no. 3 (Fall 1983), pp. 297-305.

Nevo, D. "The Conceptualization of Educational Evaluation: An Analytical Review of the Literature". Review of Educational Research 53, no. 1 (Spring 1983), pp. 117-128.

Patten, T.H. A Manager's Guide to Performance Appraisal: Pride, Prejudice, and The Law of Equal Opportunity. New York: Free Press, 1982.

Peterson, D. "Legal and Ethical Issues of Teacher Evaluation: A Research Based Approach". Educational Research Quarterly 7, no. 4 (Winter 1983), pp. 6-16.

Peterdon, D., and Ward, A. eds. Due Process in Teacher Evaluation. Washington, D.C.: University Press of America, 1980.

Peterson, K., and Kauchak, D. "Progress on Development of Lines of Evidence for the Evaluation of Public School Teaching". Paper presented at the American Educational Research Association, Montreal, April 1983 (ED 228 228).

Petrie, T.A. "Ideas That Hinder Evaluation - Debunking the Myths". NASSP Bulletin 66 (December 1982), pp. 52-55.

Phelan, W.T. "Governing Staff Reductions: The Use and Abuse of Teacher Evaluations". Education and Urban Society 15, no. 2 (February 1983), pp. 189-198.

Raths, J., and Preskill, H. "Research Synthesis on Summative Evaluation of Teaching". Educational Leadership 39, no. 4 (January 1982), pp. 310-13.

Redfern, G.B. Evaluating Teachers and Administrators: A Performance Objectives Approach. Boulder, Colorado: Westview Press, 1980.

Renihan, F.I. "The Evaluation of Principals". Saskatchewan Education Administrator (December 1980), pp. 22-31.

Riley, B.E. "Essential Phases of School Personnel Management: Communication, Conferencing, Observation and Evaluation", 1982 (ED 221 543).

Rist, M.C. "Our nationwide poll: most teachers endorse the merit pay concept". American School Board Journal 170, no. 9 (September 1983), pp. 26-33.

Robinson, T.B. Jr. "A New Approach to Increasing Organizational Commitment and Productivity - The Management Planning and Performance Appraisal System". Training and Development Journal 37, no. 1 (January 1983), pp. 58-60.

Sanzotta, D. "Assessing Your Appraisal Style". Supervisory Management 27 (December 1982), pp. 2-7.

Sapone, C.V. "Appraisal and Evaluation Systems: Perceptions of Administrators, Teachers". NASSP Bulletin 65 (1981), pp. 25-30.

Sashkin, M. Assessing Performance Appraisal. Burlington, Ontario: University Associations, 1983.

Savage, J.G. "Teacher Evaluation Without Classroom Observation". NASSP Bulletin 66 (December 1982), pp. 41-45.

Schimel, R.M. "Making Performance Evaluation Work Better: A Supervisor's Viewpoint". Public Personnel Management 11, no. 4 (Winter 1982), pp. 335-337.

Schlechty, P.C., and Vance, V.S. "Recruitment, Selection, and Retention: The Shape of the Teaching Force". The Elementary School Journal 83, no. 4 (1983), pp. 469-487.

Scontrino, M.P., and Fleenor, C.P. "Performance Appraisal and Productivity: A Natural Link". Journal of Contemporary Business 11, no. 2 (1982), pp. 832.

Seidman, W.H. "Goal Ambiguity and Organizational Decoupling: The Failure of Rational Systems Program Implementation". Educational Evaluation and Policy Analysis 5, no. 4 (Winter 1983), pp. 399-413.

Silverman, B.R.S. "Why the Merit Pay System Failed in the Federal Government". Personnel Journal 62, no. 4 (April 1983), pp. 294-297.

Smith, A., ed. Evaluating Faculty and Staff. San Fransisco: Jossey-Bass, Inc, 1983.

Solomon, G. "Teachers Rating Principals: A Sobering Experience". Principal 62, no. 4 (March 1983), pp. 14-17.

Starling, G. "Performance Appraisal in the Z Organization". Public Personnel Management 11, no. 4 (Winter 1982), pp. 343-351.

Stengel, R. "Quality Not Just Quantity: The Paideia Proposal Aims to Reform America's Schools". Time, September 6, 1982, p. 87.

Stow, S.B., and Manatt, R.P. "Administrator Evaluation Tailored to Your District or Independent School". Educational Leadership 39, no. 5 (February 1982), pp. 353-357.

Strike, K.A., and Millman, J. "Nontechnical Questions About Teachers Evaluation Systems in Elementary and Secondary Schools: A Research Agenda". Educational Evaluation and Policy Analysis 5, no. 4 (Winter 1983), pp. 389-397.

Sweeney, J. "Research Synthesis on Effective School Leadership". Educational Leadership 39, no. 5 (February 1982), pp. 346-52.

Sweeney, J. "The Post-Observation Conference: Key to Teacher Improvement". The High School Journal 66, no. 2 (December-January 1983), pp. 135-140.

"Teacher Evaluation: A CTF Conference Report". Report of a Workshop Sponsored by the Canadian Teachers' Federation, September 1982. (ED 224 105).

Thompson, F.J. "The Performance Appraisal of Public Managers: Inspiration, Consensual Tests and the Margins". Public Personnel Management 11, no. 4 (Winter 1982), pp. 306-313.

Toll, S. "This merit pay plan for management improves performance and sets an example for other employees (like teachers, for instance)". American School Board Journal 170, no. 8 (August 1983), pp. 38-9.

Totusek, P.F., and Christner, C.A. "So You Think You Want to Develop an Administrator Evaluation System". Paper presented at the Annual Meeting of the Amerjcan Educational Research Association, April 1983. (ED 228 294).

Walters, D.L. "The Measurement of Administrative Competencies". Phi Delta Kappan (February 1980), pp. 423-425.

Weick, K.E. "Administered Education in Loosely Coupled Schools". Phi Delta Kappan 63, no. 10 (1982), pp. 673-76.

Weick, K.E. "Educational Organizations as Loosely Coupled Systems". Administrative Science Quarterly 21 (March 1976), pp. 1-19.

Wells, R.G. "Guidelines For Effective and Defensible Performance Appraisal Systems". Personnel Journal (October 1982), pp. 776-782.

Whitbeck, P.H. "A Critique on the Theme of Change in Performance Appraisal". Public Personnel Management 11, no. 4 (Winter 1982), pp. 340-342.

Wuhs, S.K., and Manatt, R.P. "The pace of mandated teacher evaluation picks up". American School Board Journal 170, no. 5 (May 1983), p. 28.

Zammuto, R.F.; London, M.; and Rowland, K.M. "Organization and Rater Differences in Performance Appraisals". Personnel Psychology 35, no. 3 (Autumn 1982), pp. 643-658.

Zedeck, S., and Cascio, W.F. "Performance Appraisal Decisions as a Function of Training and Purpose of the Appraisal". Journal of Applied Psychology 67 (December 1982), pp. 752-758.